

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 06:23:18 ; Search time 11588 Seconds
(without alignments)
11436.401 Million cell updates/sec

Title: US-10-624-201a-1

Perfect score: 2735

Sequence: 1 catcgagagataaaatata.....gcaaaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2735	100.0	2735	8	AF406697 Solanum t
2	2135.8	78.1	2687	8	BT013459 Lycopersi
3	2097.8	76.7	2636	8	AF375966 Lycopersi
4	524.6	19.2	1898	8	AF406698 Solanum t
5	401.2	14.7	2128	8	AF406702 Solanum t
6	324.6	11.9	2043	6	AX506642 Sequence
7	324.6	11.9	2043	8	BT008422 Arabidops
8	324.6	11.9	2074	8	AY096548 Arabidops
9	324.6	11.9	2385	8	AF353094 Arabidops
10	324.6	11.9	2572	8	AY072175 Arabidops
11	324.6	11.9	2578	8	AY099854 Arabidops
12	295.6	10.8	1896	8	AK104872 Oryza sat
13	295.6	10.8	2476	8	AK121356 Oryza sat
14	295.6	10.8	3030	8	AK070465 Oryza sat
15	295.2	10.8	2065	8	AF406703 Solanum t
16	278.8	10.2	2523	8	AY524972 Oryza sat
17	278.8	10.2	2570	8	AK067939 Oryza sat
18	278.8	10.2	2630	8	AK069994 Oryza sat
19	266.4	9.7	2952	8	AK070543 Oryza sat

20	259.4	9.5	1449	6	AX412914 Sequence
21	259.4	9.5	1449	8	BT000609 Arabidops
22	259.4	9.5	2126	8	AY056796 Arabidops
23	250.6	9.2	1731	8	AF406700 Solanum t
24	249	9.1	1907	8	AF375964 Lycopersi
25	241.8	8.8	1939	8	AF406701 Solanum t
26	238	8.7	1840	8	AF406699 Solanum t
27	234.4	8.6	1634	8	BT000133 Arabidops
28	234.4	8.6	1824	8	AY085278 Arabidops
29	234.4	8.6	1946	8	AY598452 Arabidops
30	234.4	8.6	1950	8	AF353093 Arabidops
31	234.4	8.6	1973	8	AY062545 Arabidops
32	233.4	8.5	3037	8	AF053769 Malus x d
33	230.2	8.4	1906	8	AY570508 Arabidops
34	229.8	8.4	3132	8	GN318871 Gnetum gn
35	225.4	8.2	1599	6	AX507459 Sequence
36	225.4	8.2	1599	8	BT012291 Arabidops
37	224	8.2	1884	6	AX506210 Sequence
38	224	8.2	1884	8	AY142019 Arabidops
39	224	8.2	2050	8	AF353092 Arabidops
40	224	8.2	2124	8	AY058179 Arabidops
41	222.6	8.1	2545	8	AF173816 Arabidops
42	222.6	8.1	3162	8	AY050459 Arabidops
43	219.4	8.0	28712	8	AC145458 Cicer ari
44	204.8	7.5	2636	8	AK068034 Oryza sat
45	204.8	7.5	2698	8	AK099829 Oryza sat

ALIGNMENTS

RESULT 1
AF406697
LOCUS
DEFINITION Solanum tuberosum BEL1-related homeotic protein 5 (Bel5) mRNA, linear PLN 14-JUL-2003
complete cds.
ACCESSION AF406697.1 GI:22652114
VERSION
KEYWORDS
SOURCE Solanum tuberosum (potato)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.
TITLE Interacting transcription factors from the three-amino acid loop extension superclass regulate tuber formation
JOURNAL Plant Physiol. 132 (3), 1391-1404 (2003)
PUBMED 12857821
REFERENCE
AUTHORS Chen,H., Rosin,F.M. and Hannapel,D.J.
TITLE A KNOX protein of potato interacts with several members of the TALE family of transcription factors
JOURNAL Unpublished
AUTHORS Chen,H., Rosin,F.M. and Hannapel,D.J.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Department of Horticulture, Iowa State University, 257 Horticulture Hall, Ames, IA 50011-1100, USA
FEATURES
source
1..2735
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Desiree"
/db_xref="taxon:4113"
1..2735
/gene="Bel5"
148..2214
/gene="Bel5"
/note="StBel5"
/codon_start=1
/product="BEL1-related homeotic protein 5"

/protein_id="AA003621.1"
/db_xref="GI:22652115"
/translation="MYQTSNTNINQADHQORHNHNSNNNINOTLYLMPNNYMQG
YITSDTQQQQQLFLNSPAAALCHANIQAPQQQHFVGVPLFVSLHQIQNHG
LIORMNNQDQOQVIVPSTGVSATSCGGITTDLASQLAFORPIPTPHRQQQQQ
GLSLSPQLQQQISFNNNISSSSPRTNNVTIRGLDSSNNVLGSKYLKAAQELL
EVVNIYVKGSIKGDQDKDNMKNESPLASDVNTNSGGGESSRSKQNEVAVELTAAQ
ROELQMKAKLAMLLEEVORVROVHHOMQIIVLSEFVAGIGSAKSYTOLALHAISK
QPRCLDATAEQVKATSKLGBEELGGKLEGRSLKFDVHHLRQORALQOIQIMQPN
WPRQGLPERAVSVLRKWLFEHLFHPYKDSKIMLAKOTGLTRSOVSNWFINARVRL
WKPMEVMEYLEVRQEQNSTNSGDNKNKETNISAPNEEKPIITSSLLQDITTTQ
AISTSTISTPTAGASLHHAHNFSLGFMNDNTTTVDHIEENNAKQORNDMHKFS
SILSSVDEAKARESNKGFNPLMAAYMGDFGRFDPHQOQMTAFPHGNVSLTL
GLPPSNLAMPVSYQNYLNSDLGRSEMGSHYRMGYENIDFQSGNKRFPPTQLLPDFV
TGNIGT"

ORIGIN

Query Match		100.0%;	Score 2735;	DB 8;	Length 2735;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2735;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CATGCAGAGATAAAATATAGATCAGTCTGCACAAGAGGCAACTTCTCAAAGCTTAGAGA	60		
DB	1	CATGCAGAGATAAAATATAGATCAGTCTGCACAAGAGGCAACTTCTCAAAGCTTAGAGA	60		
QY	61	GCTTACCACCGAAGATAGACAGTTAGTTACATGCTACTGTTATAGATAAAAGGAGAAATCC	120		
DB	61	GCTTACCACCGAAGATAGACAGTTAGTTACATGCTACTGTTATAGATAAAAGGAGAAATCC	120		
QY	121	GAAGAAGAAGAAATTTTTTTTGCAGATATGTACTATCAAGGAACCTCGGATAATACTAAT	180		
DB	121	GAAGAAGAAGAAATTTTTTTTGCAGATATGTACTATCAAGGAACCTCGGATAATACTAAT	180		
QY	181	ATCAAGCTGATCATCAACAAAGCTCATATCATGGAATAGTAAATTAATTAATTTACAG	240		
DB	181	ATCAAGCTGATCATCAACAAAGCTCATATCATGGAATAGTAAATTAATTAATTTACAG	240		
QY	241	ACACTTTATTTGATGAACCTCAACATATATGCAAGGCTACACTACTTCTGCACACACAG	300		
DB	241	ACACTTTATTTGATGAACCTCAACATATATGCAAGGCTACACTACTTCTGCACACACAG	300		
QY	301	CAGCAGCAGCAGTTACTTTTCTCGAAATCTTCAACGACGACGCAACGCGTTTGCCTAT	360		
DB	301	CAGCAGCAGCAGTTACTTTTCTCGTAAATCTTCAACGACGACGCAACGCGTTTGCCTAT	360		
QY	361	GCGAATATACACACGCGCGCTGCAACAGCAGCAGCTTTGTGGTGTGCTCTTCCGGCA	420		
DB	361	GCGAATATACACACGCGCGCTGCAACAGCAGCAGCTTTGTGGTGTGCTCTTCCGGCA	420		
QY	421	GTAAGTTTGCAGCATCAGATCAATCATCATGAGCTTTTACAGGCGCATGTGGAACACCAA	480		
DB	421	GTAAAGTTTGCAGCATCAGATCAATCATCATGAGCTTTTACAGGCGCATGTGGAACACCAA	480		
QY	481	GATCAATCTCAGCAGGTAGTACCATCGTCGACGGGGGTTTCTGCCACGTCATGTGC	540		
DB	481	GATCAATCTCAGCAGGTAGTACCATCGTCGACGGGGGTTTCTGCCACGTCATGTGC	540		
QY	541	GGGATCACCAACGACTTGGGGTCTCAATTTGGGGTTTTCAGAGGCGGATTCGGACACCAA	600		
DB	541	GGGATCACCAACGACTTGGGGTCTCAATTTGGGGTTTTCAGAGGCGGATTCGGACACCAA	600		
QY	601	CACCGACAGCAGCAACAAACGACGAGCGGTCTATCTTAAGCGCTTTCTCTCAGCTACAA	660		
DB	601	CACCGACAGCAGCAACAAACGACGAGCGGTCTATCTCTAAGCGCTTTCTCTCAGCTACAA	660		
QY	661	CAGCAATAGTTTCAATAACAAATTTTCACTCTCATCAACGAGCAAAATATGTACT	720		
DB	661	CAGCAATAGTTTCAATAACAAATTTTCACTCTCATCAACGAGCAAAATATGTACT	720		
QY	721	ATTAGGGGACATTAGTGGAGGCTTACGCAATGTTTTAGGCTCTAAGTATCTGAA	780		
DB	721	ATTAGGGGACATTAGTGGAGGCTTACGCAATGTTTTAGGCTCTAAGTATCTGAA	780		

QY	781	GCTGCACAAGAGCTCTTCTTGATGAAGTGTGTTAATATTTGTTGAAAAAGCATCAAAGGAGAT	840
DB	781	GCTGCACAAGAGCTCTTCTTGATGAAGTGTGTTAATATTTGTTGAAAAAGCATCAAAGGAGAT	840
QY	841	GATCAAAAAAGAGATAAATTTCAATGAATTAAGAAATCAATGCTTTTGGCTAGTGTCAAC	900
DB	841	GATCAAAAAAGAGATAAATTTCAATGAATTAAGAAATCAATGCTTTTGGCTAGTGTCAAC	900
QY	901	ACTAATAGTTCTCGTGGTGGTGAAGTAGCAGCAGCGCAGAAAAATGAAAGTCTCTGTGAG	960
DB	901	ACTAATAGTTCTCGTGGTGGTGAAGTAGCAGCAGCGCAGAAAAATGAAAGTCTCTGTGAG	960
QY	961	CTTACAACCTGCTCAAAGACAAGAACTTCAAAATGAAAAAGCAGCTTCTTGCCATGCTT	1020
DB	961	CTTACAACCTGCTCAAAGACAAGAACTTCAAAATGAAAAAGCAGCTTCTTGCCATGCTT	1020
QY	1021	GAAGAGGTGGACAAAGGTACAGACGATACCATCAACCAATGCAAAATTAATTTATATCA	1080
DB	1021	GAAGAGGTGGACAAAGGTACAGACGATACCATCAACCAATGCAAAATTAATTTATATCA	1080
QY	1081	TTTGTAGCAAGTAGCAGGAATTTGGATCAGCCAAATCATACACTCAATTTAGCTTTGCATGCA	1140
DB	1081	TTTGTAGCAAGTAGCAGGAATTTGGATCAGCCAAATCATACACTCAATTTAGCTTTGCATGCA	1140
QY	1141	ATTTCCGAAGCAATTCAGATGCCCTAAAGGATGCAATTTGCTGAGCAAGTAAGGCGCAGCAGC	1200
DB	1141	ATTTCCGAAGCAATTCAGATGCCCTAAAGGATGCAATTTGCTGAGCAAGTAAGGCGCAGCAGC	1200
QY	1201	AAGAGTTTAGTTGAAGAGGAGCTTTGGAGGGAAAAATCGAAGGCTCAAGACTCAAAATTT	1260
DB	1201	AAGAGTTTAGTTGAAGAGGAGCTTTGGAGGGAAAAATCGAAGGCTCAAGACTCAAAATTT	1260
QY	1261	GTGGACCATCATCTAAGGCAACAAACGCGCGCTGCAACAGATAGGAATGATCAACCAAT	1320
DB	1261	GTGGACCATCATCTAAGGCAACAAACGCGCGCTGCAACAGATAGGAATGATCAACCAAT	1320
QY	1321	GCTTGGAGACCCCAAGAGGTTTACCTGAAAGAGCTGTCTGTCTCTCGTCTGTGGCTT	1380
DB	1321	GCTTGGAGACCCCAAGAGGTTTACCTGAAAGAGCTGTCTGTCTCTCGTCTGTGGCTT	1380
QY	1381	TTTCGACATTTTCTTCATCTTACCACAAAGGATTTACCTGAAAGAGCTGTCTGTCTGTGGCTT	1440
DB	1381	TTTCGACATTTTCTTCATCTTACCACAAAGGATTTACCTGAAAGAGCTGTCTGTCTGTGGCTT	1440
QY	1441	ACGGGCTTAAACAGGAGCCAGGTGTCTACTCGTTTCAATAATGCTCGAGTTCGATTATCG	1500
DB	1441	ACGGGCTTAAACAGGAGCCAGGTGTCTACTCGTTTCAATAATGCTCGAGTTCGATTATCG	1500
QY	1501	AAGCCAAATGGTAGAAGAAATGTACTTTGGAAGAGTGAAGAAATCAAGAAACAAACAGTACT	1560
DB	1501	AAGCCAAATGGTAGAAGAAATGTACTTTGGAAGAGTGAAGAAATCAAGAAACAAACAGTACT	1560
QY	1561	AATACCTTCAAGAGATAAACHAAAAACAAAGACCAATATTAAGTCTCCAAATGAAGAGAA	1620
DB	1561	AATACCTTCAAGAGATAAACHAAAAACAAAGACCAATATTAAGTCTCCAAATGAAGAGAA	1620
QY	1621	CATCAAAATTTTACTAGCAGCTTATTACAAGATGGTATTACTACTCTCAAGCAGAAATTT	1680
DB	1621	CATCAAAATTTTACTAGCAGCTTATTACAAGATGGTATTACTACTCTCAAGCAGAAATTT	1680
QY	1681	TCTACCTCAACTATTTCACATTTCCCTACTGCGAGGTGCTTCACTTCATCATGCTCACAAT	1740
DB	1681	TCTACCTCAACTATTTCACATTTCCCTACTGCGAGGTGCTTCACTTCATCATGCTCACAAT	1740
QY	1741	TTCTCTCTTCTTGGTTCATTCAACATGGATATATCTACTACTCTACTCTCATATTGAA	1800
DB	1741	TTCTCTCTTCTTGGTTCATTCAACATGGATATATCTACTACTCTACTCTCATATTGAA	1800
QY	1801	AACACGCGAAAAAGCAAGAAATGACATGACAGTCTTCTCCAGTAGTATTCTTTCA	1860
DB	1801	AACACGCGAAAAAGCAAGAAATGACATGACAGTCTTCTCCAGTAGTATTCTTTCA	1860
QY	1861	TCTGTTGACATGGAAGCCCAAGCTAGAGAAATCATCAAAATAAAGGGTTTACTAATCTCTTA	1920

||||| 1861 TCTGTTGACATGGAAGCCAAAGCTAGAGAAATCATCAATAAAGGGTTTACTTAATCCCTTTA 1920
Qy 1921 ATGCGAGCATACGCGATGGGAGATTTTGGGAAGGTTTGATCTCTCATGATCAACAATGACC 1980
Db 1921 ATGCGAGCATACGCGATGGGAGATTTTGGGAAGGTTTGATCTCTCATGATCAACAATGACC 1980
Qy 1981 GCGAATTTTCATGGAATAATATGTTGTCTCTCTTACTTTAGGACTTCTCTCTCTGAAAAAC 2040
Db 1981 GCGAATTTTCATGGAATAATGTTGTCTCTCTTACTTTAGGACTTCTCTCTCTGAAAAAC 2040
Qy 2041 CTAGCCATGCCAGTGAGCCACAAATTTACCTTTCTAATGACTTTGGGAAGTAGTCTCGAA 2100
Db 2041 CTAGCCATGCCAGTGAGCCACAAATTTACCTTTCTAATGACTTTGGGAAGTAGTCTCGAA 2100
Qy 2101 ATGGGGAGTCATTACAAATAGGAATGGGATATGAAACCAATTTGATTTTCAGAGTGGGAATAAG 2160
Db 2101 ATGGGGAGTCATTACAAATAGGAATGGGATATGAAACCAATTTGATTTTCAGAGTGGGAATAAG 2160
Qy 2161 CGATTTCCGACTCAACTATTACAGATTTTGTACAGGTAACTTAGGAACATGAATACCA 2220
Db 2161 CGATTTCCGACTCAACTATTACAGATTTTGTACAGGTAACTTAGGAACATGAATACCA 2220
Qy 2221 GAAAGTCTCGTATTGATAGCTGAAAGATAAAGGAAGTTAGGATCTCTTATTTGTG 2280
Db 2221 GAAAGTCTCGTATTGATAGCTGAAAGATAAAGGAAGTTAGGATCTCTTATTTGTG 2280
Qy 2281 TGAGGCCCTTCGGCCCAAGTCGGAGGACCCCAATTTGATACAACCTATCATAGGAGAAAAAG 2340
Db 2281 TGAGGCCCTTCGGCCCAAGTCGGAGGACCCCAATTTGATACAACCTATCATAGGAGAAAAAG 2340
Qy 2341 AAGTGGAGACTAAATTTAAAGTAAACAAATTTTAAAGCACACTTTCTAGTATATATATCTTC 2400
Db 2341 AAGTGGAGACTAAATTTAAAGTAAACAAATTTTAAAGCACACTTTCTAGTATATATATCTTC 2400
Qy 2401 TTTTATTTATAGTATAGGAAGAGAGATTTTGTGCTTTAGTGTATAGATAGTCTACT 2460
Db 2401 TTTTATTTATAGTATAGGAAGAGAGATTTTGTGCTTTAGTGTATAGATAGTCTACT 2460
Qy 2461 TAGTATAGTTATATCTTCTAGTTCCTTGAGAGAGATTGATACAACCTAGTAGTATTTTGTGCT 2520
Db 2461 TAGTATAGTTATATCTTCTAGTTCCTTGAGAGAGATTGATACAACCTAGTAGTATTTTGTGCT 2520
Qy 2521 CTTTGTGGTGGCTGGAGTACTATTTTAAAGTAACTAGTTTAAAGTAACTAGTTTAAAGTAACTAG 2580
Db 2521 CTTTGTGGTGGCTGGAGTACTATTTTAAAGTAACTAGTTTAAAGTAACTAGTTTAAAGTAACTAG 2580
Qy 2581 TAAAGTTGTGATATTTGTTCTCTCAATTTGCATATATTTGAAATATTTTGTACCTACTA 2640
Db 2581 TAAAGTTGTGATATTTGTTCTCTCAATTTGCATATATTTGAAATATTTTGTACCTACTA 2640
Qy 2641 GCTAGTCTCTAAATTTATGTTTCCATTTGCTTGAATTTGCAATTTTATTTGAAATTTTGTGCT 2700
Db 2641 GCTAGTCTCTAAATTTATGTTTCCATTTGCTTGAATTTGCAATTTTATTTGAAATTTTGTGCT 2700
Qy 2701 ATCATTATTAGATTAGCAAAAAAATTTTAAAGTAACTAGTTTAAAGTAACTAGTTTAAAGTAACTAG 2735
Db 2701 ATCATTATTAGATTAGCAAAAAAATTTTAAAGTAACTAGTTTAAAGTAACTAGTTTAAAGTAACTAG 2735

RESULT 2
BT013459
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 2687)

Kirkness, E.P., Wang, W. and Vazeille, A.
Direct Submission
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
source
1. 2687
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132117P"
/tissue_type="fruit"
/note="TZABC46"
ORIGIN
Query Match 78.1%; Score 2135.8; DB 8; Length 2687;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 2452; Conservative 0; Mismatches 132; Indels 107; Gaps 13;
Qy 12 AAAAAATATAGATCAGTCTGACAAGAGGCAACTTCTCAAGAGCTTAGAGAGCTACCCCG 71
Db 25 AAAAAATATAGATCAGTCTGACAAGAGGCAACTTCTCAAGAGCTTAGAGAGCTACCCCG 84
Qy 72 AAGATAGACAGTTAGTTTACA--TGTAAGTTATAGATAAAGGGAATCCGAGAGAA 129
Db 85 AAGATAGACAGTTAGTTTACATGTGTAATTTGATAAAGGAGAAATTTGAACAAG-- 142
Qy 130 AGAATTTTTTTTTCAGATATGTACTATCAAGGAACCTCGGATATATCTAATATACAAGCT 189
Db 143 TTTTGTGAGATATGTACTATCAAGGAACCTCGGATATATCTAATATACAAGCTGAT 195
Qy 190 GATCATCAACAACGTCATATATGGAATAGTAATTAATTAATTAATTAATTAATTAATTAAT 249
Db 196 CATCATCAACAACAATTAATCTTGGGAATAGTAATTAATTAATTAATTAATTAATTAATTAAT 255
Qy 250 TTGATGAACCTTAACCAATTAATGCAAGGCTACACTTCTTCAACAGCAGCAAGCAAC 303
Db 256 CTGATGAACCTTAACCAATTAATGCAAGGCTACACTTCTTCAACAGCAGCAAGCAAC 315
Qy 304 CAGCAGCAGTTTACTTTTCTGAAATTTCTTCAACAGCAGCAAGCAAC 348
Db 316 CAACAGCAACAGAACCCAGCAGTCTTCTGAAATTTCAACAGCAGCAAGCAAC 375
Qy 349 GCGCTTTGCCATCGGAATATACAAACGCGCGCTGCAACAGCAGCACTTTTGTGCGGTG 408
Db 376 GCGCTTTGCCATCGGAATATACAAACGCGCGCTGCAACAGCAGCACTTTTGTGCGGTG 435
Qy 409 CTTCTTTCCGCGAGTTAGTTTGCAGATCAGATCAATCATCATCTGACACTTTTACAGCGCATG 468
Db 436 CCGCTTTCCGCGAGTTAGTTTGCATGATCAGATCAATCATCATCTGACACTTTTACAGCGCATG 495
Qy 469 TGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCATCGTACGCGGGGTTTCTGCC 528
Db 496 TGAACAACATCAAGACCAATCTCAGCAGGTGATAGTACCGTCTGACGCGGTGTTCTGCC 555
Qy 529 ACGTATGCGCGGATCACCAAGGCTTGGCGTCTCAATTTGGCGTTTACAGAGCCGAT- 587
Db 556 ACGTATGCGCGGATCACCAAGGCTTGGCGTCTCAATTTGGCGTTTACAGAGCCGAT 615
Qy 588 TCCGACACCAACCAACCCAGCAGCAACCAACAGCAGCAAGGCGGTCTATCTCTA 639
Db 616 GTGGTGTCTCCGACACCAACCAACCCAGCAGCAACCAACAGCAGCAAGGCGGTCTATCACTG 675
Qy 640 AGCCTTTTCTCCTCAGCTACCAAGCAAAATTTAGTTTCAATAACAATATTTTCATCTCATCA 699
Db 676 AGCCTTTTCTCCTCAGCAACAGCAAAATTTAGTTTCAATAACAATATTTTCATCATCATCA 735
Qy 700 CCAAGGACAAATTAATGTTTACTTATAGGGGAACATTAAGTGAAGTCTTAGCAACATGGTT 759
Db 736 CCAAGGACAAATTAATGTTTACTTATTAGGGAACAAATGGAATGATGTTTCTAGCAACATGAT 795
Qy 760 TTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTCAAGTGTGTTATATATGTT 819
Db 796 TTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTCAAGTGTGTTATATATGTT 855

/mol_type="mRNA"		/db_xref="taxon:4081"	
1. .2636			
/gene="BL2"			
100. .2199			
/gene="BL2"			
/note="TBL2"			
/codon_start=1			
/product="bell-like homeodomain protein 2"			
/protein_id="AAP47025.1"			
/db_xref="GI:31323447"			
translation="MYVQGTSDNNIQAHHQOHHNNLGNSSNNIQTLLYLANPNSYMGQ YTTDTQHLQQOQOQHLLFLNSAPAGGNALSHANIQHAPLOQHFVGLPAPVSLH DQINHHGLQMNNDQSQVIVPSSVTSVATSCGGTTTDLASQLAFQRPVIVSPTP OHRQOQOQGLSLSPQOQOQISFNNNTSSSSPRTNNVTIRGTMDGCSNMLGSK YLKAAQELDELIVNGSKNGDDQKDKNSMNELIPLVDVNTNSGGGGSSRSRQ KNEVAILTTAOROLEQMKAKLAMLEEVORVROVHQMOLIVSSFEQVAGVGSNAK SYTOLALHATSKQFRLKDAISEQVKATSKSLGDEGLGKIEGSRLLKFDVHHLRQOR ALQQLGMQPNWAPQRPGLPERAVSVLRAMLFEHLPHYPKDSKIMLAKQTGLTRSQ VSNQFINARVRLKPMVEEMYLEEVKNQOQNSNTSGDNKNKNETNISAPNEEKOPIT SSLDQGTQAEISTISPTAGASLHLAHNFSFLGSFNMENTTTVDHIENNAKK PRNHDMKFPSPSSILSSVMEAKARESTNKGFTNPLMAAVAMGDGFRFDPHDOQMTAN PHGNNGVSLTLGLPSPENLAMPVVSQQNLSNELGSRPEIGSHYNRMGEVNIQFQSGNK RPTQLLPDPFTVGNLGT"			
ORIGIN			
Query Match 76.7%; Score 2097.8; DB 8; Length 2636;			
Best Local Similarity 90.9%; Pred. No. 0;			
Matches 2418; Conservative 0; Mismatches 132; Indels 111; Gaps 13;			
Qy	46	CTCAAGCTTAGAGAGCTACACCGGAAGATAGACAGTTAGTTACAT-----GTACTGT	99
Db	1	CTCAAGCTTAGAGAGCTACACCGGAAGATAGACAGTTAGTTACATGTGTGAGTACTGT	60
Qy	100	TATAGATAAAGGAGAAATCGAAGAGAGAAAGAAATTTTTTTCAGAGATATGCTACTCAA	159
Db	61	TATTTGATAAAGGAGAAATTTGAACAAG-----TTTTGCGAGATATGCTACTCAA	111
Qy	160	GGAACTCGGATAATATCAAGCTGATCATCAACACGCTCATATATCATGGAAT	219
Db	112	GGAACTCGGATAATATATCAAGCTGATCATCAACACGCTCATATATCAATATCTTGGG	171
Qy	220	AGTAATATAATATATTCAGACACTTTATTTGATGAACCCCTAACATATATGCAAGGC	279
Db	172	AATAGTATAATATAATTCAGACGCTGATCTGATGAACCCCTAACAGTTATATGCAAGGC	231
Qy	280	TACACTACTCTGACACACAGCAG-----CAGCAGCAGTTACTT	318
Db	232	TACACTCAACTGACACACAGCAGCAGCTGATCAACACAGCAACGAACCCAGCACCGTGT	291
Qy	319	TTCTTGAATTTCTTACCAGCAGCAAGCAACGCGCTTTTGCCATGCGAATATACACACGCG	378
Db	292	TTTCTGAATTTCCAGCACCAGCAGGAGGCAACGCGCTTAGCCATGCGAATATACACACGCG	351
Qy	379	CCGCTGCAACAGCAGCACTTTGTCGGTGTGCTCTTCGGCAGTAAAGTTTGCACGATCAG	438
Db	352	CCTCTGCAACAGCAGCACTTTGTCGGTGTGCTCTTCGGCAGTAAAGTTTGCATGATCAG	411
Qy	439	ATCAATCATCATGACATTTTACAGCGATGTTGAACCAACCAAGATCAATCTTCAGCAGGTG	498
Db	412	ATCAATCATCAGGACATTTTACAGCGATGTTGAACCAATCAAGACCAATCTTCAGCAGGTG	471
Qy	499	ATAGTACCATCGTCGACGGGGTTTCTGCCACGTCATGTGGCGGGATCACCCAGGACTTG	558
Db	472	ATAGTACCGTCGTCGACGGTGTCTTCTGCCACGTCATGTGGCGGGATCACCCAGGACTTG	531
Qy	559	GCCTCTCAATTTGGGGTTTTCAGAGCGGAT-----TCCGACACCAACACCCGACAG	609
Db	532	GCCTCTCAACTGGGGTTTTCAGAGCGGATTTGTGGTGTCTCCGACACCAACACCCGACAA	591
Qy	610	CAGCAACACAGCAGCGGGTCTATCTCTAGCCCTTTCTCTCAGCTACACAGCAATTT	669
Db	592	CAACACACAGCAGCGGGTCTATCTAGCCCTTTCTCTCAGCAACAGCAACAAATTT	651
Qy	670	AGTTTCAATAACCAATATTTTCATCTCTCATCACCAGGACAAATAATGTTTACTATTAGGGGA	729
Db	652	AGTTTCAATAACCAATATTTTCATCATCATCATCACCAGGACAAATAATGTTTACTATTAGAGGA	711
Qy	730	ACATTAGATGGAAGTTCTTAGCAACATGGTTTATAGGCTCTAAGATATCTGAAAGCTGCACAA	789
Db	712	ACAATGGATGGATGTTCTTAGCAACATGATTTTATAGGCTCTAAGATATCTGAAAGCTGCACAA	771
Qy	790	GAGCTTCTCATGAAGTTGTTAATATTGTTGGAAAAAGCATCAAGAGAGATGATCAAAAAG	849
Db	772	GAGCTTCTCATGAAGTTGTTAATATTGTTGGAAAAAGTAAACAAAGAGATGATCAAAAAG	831
Qy	850	AAGGATAATTTCAATGAATAAAGAAATCAATGCCTTTGGCTAGTGTCAACACATAATAGT	909
Db	832	AAGGATAATTTCAATGAATAAAGAAATTAATTCCTTTGGTTAGTGTCAATACTAATAGT	891
Qy	910	TC-----TGGTGGTGGTGAAGTAGCAGCAGGCGAGCAAAAAAATGAAGTTGCTGTGAGCTT	963
Db	892	TCTGGTGGTGGTGGTGAAGTAGCAGCAGGCGAGCAAAAAAATGAAGTTGCTGTGAGCTT	951
Qy	964	ACAACTGCTCAAGACAAAGAACTTCAAAATGAAAAAGCCCAAGCTTCTTGGCCATCTTGAA	1023
Db	952	ACAACTGCTCAAGACAAAGAACTTCAAAATGAAAAAGCCCAAGCTTCTTGGCCATCTTGAA	1011
Qy	1024	GAGTGGAGCAAAAGGTTACAGACAGTACCATCACCAAAATGCAAAATAATTTGTTATTATCATTT	1083
Db	1012	GAGTGGAGCAAAAGGTTACAGACAGTACCATCACCAAAATGCAAAATAATTTGTTATTATCATTT	1071
Qy	1084	GAGCAAGTAGCAGGAATTTGGATCAGCCAAATCATACACTCAATTTAGCTTTGTCATGCAATTT	1143
Db	1072	GAGCAAGTAGCAGGAATTTGGATCAGCAAAATCATACACTCAATTTAGCTTTGTCATGCAATTT	1131
Qy	1144	TCGAAGCAATTTAGATGCTTAAAGGATGCAATTTAGTGAAGGCGACGAGCAAG	1203
Db	1132	TCGAAGCAATTTAGATGCTTAAAGGATGCAATTTAGTGAAGGCGACGAGCAAG	1191
Qy	1204	AGTTTAGGTGAAGAGGAGCTTTGGGAGGAAAAATCGAAGCTCAAGACTCAAAATTTGTG	1263
Db	1192	AGTTTAGGTGAAGAGGAGCTTTGGGAGGAAAAATCGAAGCTCAAAATTTGTG	1251
Qy	1264	GACCATCATCTAAGGCAACAAACGCGCGCTGCAACAGATAGGAATGATGCAACCAATGCT	1323
Db	1252	GACCATCATCTAAGGCAACAAACGCGCGCTGCAACAGATAGGAATGATGCAACCAATGCT	1311
Qy	1324	TGAGAGCCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGCTCTTCTGCTTCTGCTTTTC	1383
Db	1312	TGAGAGCCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGCTCTTCTGCTTCTGCTTTTC	1371
Qy	1384	GAGCATTTTCTTTCATCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTT	1443
Db	1372	GAGCATTTTCTTTCATCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTT	1431
Qy	1444	GGGCTAACAAGGAGCCAGGTGTCTAAGTGTTCATTAATGCTCGAGTTGATGATGAG	1503
Db	1432	GGGCTAACAAGGAGCCAGGTGTCTAAGTGTTCATTAATGCTCGAGTTGATGATGAG	1491
Qy	1504	CCAAATGTTAGAGAAATGTTACTTTGGAGAAAGTGAAGAAATCAAGAAACAAACAGTACTAAT	1563
Db	1492	CCAAATGTTAGAGAAATGTTACTTTGGAGAAAGTGAAGAAATCAAGAAACAAACAGTACTAAT	1551
Qy	1564	ACTTCAGGAGATACAAAAAACAAGAGACCAATATATAGTGTCCAAATGAAGAGAAACAT	1623
Db	1552	ACTTCAGGAGATACAAAAAACAAGAGACCAATATATAGTGTCCAAATGAAGAGAAACAA	1611
Qy	1624	CCAAATTTACTAGCAGCTTATTAAGATGGTATTAATGCTTCTTACTACTCAAGCAAAATTTCT	1683
Db	1612	CCAAATTTACTAGCAGCTTATTAAGATGGTATTAATGCTTCTTACTACTCAAGCAAAATTTCT	1665
Qy	1684	ACCTCAACTATTTCAACTTCCCTTACTGAGGCTTCTTCACTTCAATGCTCAACATTTTC	1743
Db	1666	ACCTCAACTATTTCAACTTCCCTTACTGAGGCTTCTTCACTTCAATGCTCAACATTTTC	1725

1744 TCCTTCCTTGTTGTTCAATCAACGATGATAATACTACTACTACTCTGTTGATCATATATTGAAAC 1803
Db |||||
1726 TCCTTCCTTGTTGTTCAATCAACGATGATAATACTACTACTACTCTGTTGATCATATATTGAAAC 1785
Qy |||||
1804 AACGCGAAAAAGCAAGAA---ATGACATGCAACAGTTTCTCCAAAGTAGTATTTCTTTCA 1860
Db |||||
1786 AACGCGAAAAAGCAAGAAATCATGACATGCAACAGTTTCTCCAAAGTAGTATTTCTTTCA 1845
Qy |||||
1861 TCTGTTGACATGCAAGCAAGCAAGTAGAGATCATCAATAAAGGTTTACTATCTTTA 1920
Db |||||
1846 TCTGTTGAAATGGAAGCAAGCAAGTAGAGATCAACAACAAAGGTTTACTATCTTTA 1905
Qy |||||
1921 ATGCGACATACCGGATGCGAGATTTTGGAGGTTTGGAGGTTTGCCTCATGATCAACAATGACC 1980
Db |||||
1906 ATGCGAGGTACCGGATGCGAGATTTTGGAGGTTTGGAGGTTTGCCTCATGATCAACAATGACC 1965
Qy |||||
1981 GCGAATTTTCATGGAATATGTTGTTCTCTTTAGGACTTCTCTCTCTCTGAAAC 2040
Db |||||
1966 GCGAATTTTCATGGAATATGTTGTTCTCTTTAGGACTTCTCTCTCTCTGAAAC 2025
Qy |||||
2041 CTAGCCATGCGAGTGCACCAAAATTTACCTTTCTAATGACTTGGAGTAGTCTGAA 2100
Db |||||
2026 CTAGCCATGCGAGTGCACCAAAATTTATCTTTCTAATGAGTTGGAGTAGGCTGAA 2085
Qy |||||
2101 ATGGGAGTCAATACATAGATGCGATATGAAACATTTGATTTTCAGAGTGGGAATAAG 2160
Db |||||
2086 ATAGGAAGTCAATACATAGATGCGATATGAAACATTTGATTTTCAAGTGGGAATAAG 2145
Qy |||||
2161 CGAATTTCCGACTCAACTATTACAGATTTTGTTCAGGTAATCTAGGAACATGAATACCA 2220
Db |||||
2146 CGAATTTCCAACTCAACTATTGCGAGATTTTGTTCACAGGTAATCTTGGAAATGAATACCA 2205
Qy |||||
2221 GAAAGTCTCGATTGATGCTGAAAGAT---AAAGGAGTTAGGATCTCTTATATG 2278
Db |||||
2206 CAAAGTCTCGATTGATGCTGAAACATAAAAAAGGAGTTAGGATCTCTTATATG 2265
Qy |||||
2279 TGTGAGGCTCTCTGGCCCAAGTCGAGGACCCCAATTTGTATACAACTATCATAGGAA 2338
Db |||||
2266 TGTGAGGCTCTCTGGCCCAAGTCGAGGACCCCAATTTGTATACAACTATCATAGGAG--- 2322
Qy |||||
2339 AGAAGTGGAGACTAAATTAAGTAACAAATTTTAAAGCACACTTCTCTAGTATATATCT 2398
Db |||||
2323 -----ACAAAAGTTTAAAGCACACTTCTCTAGTATATATCT 2358
Qy |||||
2399 TCTTTTTT---TTATAGTATAGAAAGAGAGATTTTGTCTTGTAGTATAGATAGT 2455
Db |||||
2359 TCTTTTTTAAATTAATAGTATAGAAAGAGAGATTTTGTCTTGTAGTATAGATAGT 2418
Qy |||||
2456 CTACTTAGTATAGTTATCTTCTAGTTCCTTGAGAGATTGATACAACTAGTAGTATTT 2515
Db |||||
2419 CTACTTAGTATAGTTATCTTCTACTTCTCCTTGAGAGATTGATACAACTAGTAGTATTT 2478
Qy |||||
2516 TTTTCTTTTGG-----GTTGGCTTGGAGTACTATTTTAAAGTTAT-----TGG 2558
Db |||||
2479 TTATTTTGGGGGTTGAAATAGCTTGTCTTGGAGTACTATTTTAAAGTTAAATGTTGAAATTT 2538
Qy |||||
2559 AAAGTACTAGCTAG---TAAATGTTTAAAGTTGTGATATTTCTCTCAATTTGATATATA 2616
Db |||||
2539 AAAGTACTAGCTAGTATATGTTTAAAGTTGTGATATTTGTTACTCTCAATTTGATATATA 2598
Qy |||||
2617 ATTTGAAATATTTTGTACCTA 2637
Db |||||
2599 ATTTGAAATATTTTGTACCTA 2619

RESULT 4
AF406698
LOCUS
DEFINITION Solanum tuberosum BEL1-related homeotic protein 11 (Bell1) mRNA,
partial cds.
ACCESSION AF406698
VERSION AF406698.1 GI:22652116
KEYWORDS

Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS 1 (bases 1 to 1898)
TITLE Interacting transcription factors from the three-amino acid loop
extension superclass regulate tuber formation
JOURNAL Plant Physiol. 132 (3), 1391-1404 (2003)
PUBMED 12857821
REFERENCE
AUTHORS 2 (bases 1 to 1898)
TITLE A KNOX protein of potato interacts with several members of the TALE
family of transcription factors
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 1898)
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Department of Horticulture, Iowa State
University, 257 Horticulture Hall, Ames, IA 50011-1100, USA
FEATURES
source
Location/Qualifiers
1..1898
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Desiree"
/db_xref="taxon:4113"
<1..1898
/gene="Bell1"
<1..1608
/gene="Bell1"
/note="StBell1"
/codon_start=1
/product="BEL1-related homeotic protein 11"
/protein_id="AA003622.1"
/db_xref="GI:22652117"
/translation="MTFRSLPLDLREISTTHQVGIILSSSLPSPGVTNNINHTRG
LGSSSFISNGMILGSKYLKVAQDLDEVVNGKIKLSDGLSGAKKKHLDNELI
SLASDDVSSSKNSGVELTTAQRQELQMKAKLVSMLEVDVQRYQHYHQMTATS
FPGTIGSSSKSYTQALHTITSQPRCKDAISGQIKDTSKLTGSEENIGGKIESKL
KEVDHLLQRQALQJGHWQTMNAPQGLPERAVSLRAWLFEHLFHPYKDSKII
LAKOTGLTRQSVNFINARVLMKPMVEEMTEVKKNNQONFEPNNNEIVGSKSI
VQEKLPITSSNIHNASPNDSISTISPTGGGSIPTQTVAGSFIRSLAMNEDD
QNNKARNEMONCSTSTILSMERIINKVQDETIKSEKFNNTOTRECYSLMTPNYT
MDQFGTRFNQNHQLATTTFHQNGHVSITLGLPNSENQHNVIQLEHYNQPTH
HPNISYENIDFQSGRYATQLLDQFVS"

Query Match 19.2%; Score 524.6; DB 8; Length 1898;
Best Local Similarity 71.3%; Pred. No. 7.5e-96;
Matches 813; Conservative 0; Mismatches 274; Indels 54; Gaps 7;
Qy 729 AACATTAGATGGAAGTTCTAGCAACATGGTTTGGCTCTAAGTATCTGAAAGCTGCACA 788
Db |||||
144 ATCATCTTTTCGATTTCTTAATGGGATGATATGGGTTCTTAAGTACCTAAAAAGTTGCACA 203
Qy 789 AGAGCTCTTGTAGTGAAGTTGTTAAATATCTCGAAAAAGCATCAAGGAGATGATCAAAA 848
Db |||||
204 AGATCTCTTGTAGTGAAGTTGTTAA---TGTGAAAAAAGCATCAAAATATCAGATGGCTT 260
Qy 849 GAAGGATATTCATGAATAAAGAATCAATGCTTTGGCTAGTGTGTCACACACTAATAG 908
Db |||||
261 AGAGAGTGGTCAAGGAGAAACACAAA-----TTGGACAATGAATTAATATCTTTTGGC 314
Qy 909 TTCTGGTGGTGAAGTAGCAGAGGAGAGAAAAAATGAAGTTGCTGTGAGCTTACCAAC 968
Db |||||
315 TAGTGATGATGTTGAAGCAGCAGC-----CAAAAAAATAGTGGTGTGTAACCTTACCAAC 368
Qy 969 TGTCTCAAGACAAAGAACCTTCAATGAAAAAGCCAGCTTCTTGCCATCTCTTGAAGAGGT 1028
Db |||||
369 AGCTCAAGACAAAGAACCTTCAATGAAAGAACCCAGCTTGTGTAGCATCTCTGTATGAGGT 428
Qy 1029 GGAGCAAGGTTACAGACAGTACCATCACCAAAATGCAAAATTAATTGTATTATCATTTGAGCA 1088

Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.

Direct Submission
Submitted (15-MAY-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Shinn, P., Bowser, L., Chan, M. M., Chang, C. M., Dale, J. M., Hauan, V. W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. I., Southwick, A., Tang, C. C., Toriumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Davis, R. W., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers
1. .2043
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="U25364"
/ecotype="Columbia"
/note="This clone is in PUNI 51"

1. .2043
/note="putative homeodomain transcription factor"
/codon_start=1
/product="At2g35940"
/protein_id="PAP37781.1"
/db_xref="GI:30725518"
/translation="MAAFHGNPEISAGSGGLQTLILMNPTTYVYVYQTDNDNNNNNSNNNTNTNTNNNNNSVFLDHPAPNASQOQFVGPLSGHBAASITRADNISVLHGYPVRYQVSLYGHQVPTHQAACTPQAQGLSLTSSQQQQQHQHOHPIHVGFGSGHDIVRGSGTSGVTGNLANLVSSKYLKAAOELLDEVNADSDHMDKASHLVPSKSGSGCNPKPVGESGAGGGAAGKPRPELTGAETQELQMKAKULMLHVEQRVRYHQMQMVISFEQAAGISAKSYSLAKTISRQFCLEAKAGLKAASLSGSDSVSGVREGSLKFDVHLLRQALQQLGMIPSNNAWPKPQGLPERAVSLRWLFPEHLFHPYKDSDKMLAKQTGLTRSQVSNWFINARVLWKPWGESMYBMEKQAKNNGMSMEKTPLDQNSDASKTSQKSPMADTYHNPNHNGDLGVTYMGQSPKRLRTSDETMOPINADPSSNEKLTKMILEEQGIRSDGYPMGNGFYQDMSRPFVDSQELMAQRYSGNNNGVSLTLGLPHCDLSLSTHHQGFQMTGHGIPIGRRVKIGTSEYGEATNGSGSTTTTAHSSAAAAAYNGNIQNKRYVAQLLPDFA"

ORIGIN

Query Match	11.9%	Score 324.6	DB 8	Length 2043
Best Local Similarity	65.2%	Pred. No. 2.5e-55		
Matches 53	Conservative 0	Mismatches 284	Indels 21	Gaps 3

QY 758 TTTTAGCTCTAAGTATCTGAAAGCTGCACAAAGAGCTTCTTGATGAAGTTGTTTAATTG 817
|||||
Db 569 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAAGAGCTTCTTGACGAAGTAGTCAACGCTG 628
|||||
QY 818 TTGAAAAAGCATCAAGAGGAGATGATCAAAAAGAGGATATTCATGATTAAGAATCAA 877
|||||
Db 629 ATTCGGATGATCAAGACGCTAAATCCCAACTATTTCATCGAAAAAGGGTAGTTGCGGAA 688
|||||
QY 878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTCTCTGTTGGTGGTGAAGTAGCAGCAGC 937
|||||
Db 689 ATGATAAACCTGTCGGAGAAATCATCGCCGCGCTGGAGGAGAGGTTCCGGTCCGGAG 748
|||||
QY 938 AGAAAAATG-----AAGTTGCTGTTGAGCTTTACACTGCTTCAAGACACAGAACTTCAA 991
|||||

749	Db	CAGAACGAGCCGGGAAACGTCGGTGGAGCTAGGCACGGCAGAGAGACAAGAATACAGA	808
992	Qy	TGAAAAAGCCAAAGCTTCTTGCCATCGCTTCAAGAGGCTGGAGCAAGAGTACAGACAGTACC	1051
809	Db	TGAAGAAAGCAAACTTAGTACATGCTTCATGAGGTGGAGCAGAGATAGACAGTACC	868
1052	Qy	ATCACAAATGCAATAAATTTGTAATTATCATTTTGGAGCAAGTACAGGAATTTGGATACGCCA	1111
869	Db	ACCAGCAGATGCAGATGGTGTATCTCTTCGTTGAGCAAGCGCAGGATAGGATCAGCGA	928
1112	Qy	AATCATACACTCAATTTAGTCTTTCATGCAATTTTCGAAGCAATTTTCAGATGCTTAAGGATG	1171
929	Db	AGTCATACACGTCGCTAGCAATTTGAAACCAATATCAAGACAGATTCCTCGTTGCTTGAAGAGG	988
1172	Qy	CAATTGCTGAGCAAGTAAAGGCGCAGCAGCAAGAGTTTTCAGTGAAGGAGGCTTG----	1227
989	Db	CGATCGCTGGTCAGATATAAAGCGGCCACAAGAGTCTTGGGAGGGAAGATTTCAGTGTCTG	1048
1228	Qy	--GGAGGGGAAATCGAAGGCTCAAGACTCAAAATTTGTGGACCAATCATCTTAAGCAACAAC	1285
1049	Db	GTCTTGGGAGGTTTGGGGGTCGAGGCTCAAGTTCGTGGACCAACCACTTGGAGCAGCAAA	1108
1286	Qy	GCGCGCTGCAACAGATAGAAATGATGCAAC-----CAATGCTTGGAGACCCCAAA	1336
1109	Db	GAGCTCTTCAACAACCTGGGAATGATTCACATCTTCCAAATATGCTTGGAGACCTCAAC	1168
1337	Qy	GAGGTTTACTCTGAACAGAGCTGCTCTGCTTCGTCCTTGGCTTTTCGAGCATTTTCTTC	1396
1169	Db	GTGGTCTCCAGAACGAGCGCTCTCAGTCTCGTGCTTGGCTCTTCGAAACACTTCTTC	1228
1397	Qy	ATCTTTTACCCAAAGGATTTCAGACAAATCATGCTTCTTAAGCAAAACGGGCTAACAGGA	1456
1229	Db	ATCCATACCTTAAGGATTTCGGAACAACACATGCTAGCTAAGCAACAGGACTCACTCGTA	1288
1457	Qy	GCCAGTGTCTAACTGGTTTCATAAATGCTCGAGTTCGATTATGCAAGCCAAATGGTAGAAG	1516
1289	Db	GCCAGTGTGCACTGGTTTATAAACCGCAGAGTTCGGTTATGSGAAACCAATGGTGGAGG	1348
1517	Qy	AAATGTACTTGGGAAGTGAAGAATCAAGACAAACA	1555
1349	Db	AGATGTACATGGAGGAATGAAGGACGAGCAAGAACA	1387
<hr/>			
RESULT 8			
AY096548			
LOCUS			
DEFINITION	AY096548 2074 bp mRNA linear PLN 18-SEP-2002		
ACCESSION	Arabidopsis thaliana putative homeodomain transcription factor		
VERSION	(AT2935940) mRNA, complete cds.		
KEYWORDS	AY096548		
SOURCE	FLI CDNA		
ORGANISM	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
	1 (bases 1 to 2074)		
	Yamada, K., Ban, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.		
TITLE	Arabidopsis Open Reading Frame (ORF) Clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2074)		
AUTHORS	Yamada, K., Ban, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.		

Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.I., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Location/Qualifiers
1. .2074
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="U18537"
/ecotype="Columbia"
/notes="This clone is in pUNI 51."
1. .2074
/gene="At2g35940"
1. .2043
/gene="At2g35940"
/codon_start=1
/evidence=experimental
/product="putative homeodomain transcription factor"
/protein_id="AA20198.1"
/db_xref="GI:20465477"
/translation="MAAYFGNPPEISAGSDGLQTLILNPTTVYVYQQDNDNNNSNNNNNTNTNTNNNSFVLSHAPQPNASQFVGLPSLGHAEASITRADNISVLHGYPVVOYSLYSHQVDPTHQACETPRAQGLSLTSLSSQQQOHHQHOPIHVFGSGHEDIRVSGSTGVTNGIANLVSKYLKAAQELLDVYNADSDMNAKSQLFSSKKGCGNDKPVGESAGAGGSGGAGAGKRPVELGTABERQIOWKAKLSNMLHEVQRYQYHQOMQVVISFEQAAGIGSAKSVTSIALKTIISQPRCLKEAIAQIKAANKSLGEEDSVSGVRFEFGSRGLKFDVHLHQRALQOLGMIQHPNSNAWRPQRLPERAVSLRAWLFEHLHPYPKDSKDKMLAKQTLGRSQVSNWFNARVRLKPVVEE MYNEEKQAKNGSMKEKPTLDOSNDSKSTNSQKSPMADTTHMPNHNGLLEGVTQMGSQRLRTSDSTMQPINAQFSSNEKLTLEERQIGRSDGYPFMGNFGQYQMDMGRFDVYSQELMAQRYSGNNNGVSLTLGLPHCDLSLSTHGHQFMQTHGPIGRVKIGETREYGPATINGSGSTTTAAHSSAAAAAYNGMNIQKRVVALLPDEVA"
2044. .2074
/gene="At2g35940"

3'UTR
ORIGIN
Query Match 11.9%; Score 324.6; DB 8; Length 2074;
Best Local Similarity 65.2%; Pred. No. 2.5e-55;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;
Qy 758 TTTTAGGCTCTAGATCTCAAGAGCTGCACAGAGCTTCTTGATGAAGTTGTTAAATTG 817
Db 569 TTTGTAGCTCAAGTACTTCAAGGAGCAGCAGAGCTTCTTGACGAGTAGTCAACGCTG 628
Qy 818 TTGMAAAACATCAAGGAGATGATCAAAAGAGGATAATTCAATGAATAAAGATCAA 877
Db 629 ATTCCGATGATGACGCTAAATCCCACTATTCTCATCGAAAAAGGGTAGTTGGGAA 688

Qy 878 TGCCTTTGGCTAGTGTCAACACATAATAGTTCGGTGGTGGTGAAGTAGCAGCAGGC 937
Db 689 ATGATAAACCTGTCGAGAAATCATCGCGCGCGCTGGAGAGAGGTTCCGGTGGCGGAG 748
Qy 938 AGAAAAATG-----AAGTTGCTGTTGAGCTTACAACTGCTCAAGACAAAGAACTTCAAA 991
Db 749 CAGAAGCAGCCGGGAAAACGTCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 808
Qy 992 TGAATAAAAGCCCAAGCTTCTTGGCATGCTTGAAGAGGTGGAGCAAGAGTACAGACAGTACC 1051
Db 809 TGAAGAAAGCAAAACTTTAGTAAATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 868
Qy 1052 ATCAACCAATGCAAAATAATTGTATTATATCAATTTGAGCAAGTAGCAGAAATTGGATCAGCCA 1111
Db 869 ACCAGCAGATGAGATGGTGAATCTCTCTGTCGAGCAAGCGCAGGATAGGATCAGCGA 928
Qy 1112 AATCATACACTCAATAGCTTTTGCATGCAATTTTGAAGCAATTCAGATGCTTAAAGGATG 1171
Db 929 AGTCATACACGCTCGCTAGCATTTGAAACCATATCAAGACAGTTCCTGCTTGAAGAGAGG 988
Qy 1172 CAATTGCTCAGCAAGTAAAGCGCAGCAGCAAGTGTAGGTGAAGAGGAGGCTTG--- 1227
Db 989 CGATCCTGTCAGATTAAGCGGCCCAAGAGTCTTGGGAGGAGAGATTAGTGTCTG 1048
Qy 1228 --GGAGGGAAAAATCGAAGGCTCAAGCTCAAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
Db 1049 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACTTTGAGACAGCAAA 1108
Qy 1286 GCGCGCTGCAACAGATAGGAATGATGCAAC-----CAATGCTTGGAGACCCCAAA 1336
Db 1109 GAGCTCTTCAACAACACTGGGAATGATTCACATCTCTTCCAATATGCTTGGAGACCTCAAC 1168
Qy 1337 GAGTTTACCTGAAAGAGCTGCTCTGCTCTGCTTGGCTTTGGCTTTTCGAGCATTTTCTTC 1396
Db 1169 GTGCTCTCCAGAACGAGCGCTCTAGTTCGTGCTTGGCTTCTTCGACACTTTTCTTC 1228
Qy 1397 ATCTTTACCCAAAGGATTCAGACAAATCATGCTTGTAAAGCAAAACGGGCTTAACAGGA 1456
Db 1229 ATCCATACCTTAAGGATTCGGAACGACATGCTAGCTAAGCAAAACAGGACTCACTCGTA 1288
Qy 1457 GCGAGTGCTTAACCTGTTTCAATGCTGAGTTCGATTTGGAAGCCAAATGGTAGAAG 1516
Db 1289 GCGAGTGCTGAACCTGTTTATAAAGCGGAGAGTTCGGTTATGGAACCAATGGTGGAGG 1348
Qy 1517 AATGTTACTTGAAGAGTGAAGATCAAGAAACAAAAACA 1555
Db 1349 AGATGTACATGAGGAATGAAGGACGACGCAAGAAACA 1387

RESULT 9
AF353094
LOCUS
DEFINITION
Arabidopsis thaliana 2385 bp mRNA linear PLN 30-APR-2001
cvs.
AF353094
VERSION
AF353094.1 GI:13877516
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 2385)
Pidkowich, M.S., Samach, A., Modrusan, Z. and Haughn, G.W.
A family of BEL1-like homeodomain (BLH) proteins in Arabidopsis thaliana
Unpublished
2 (bases 1 to 2385)
Pidkowich, M.S., Samach, A., Modrusan, Z. and Haughn, G.W.
Direct Submission
Submitted (26-FEB-2001) Botany, University of British Columbia, #3529 - 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
Location/Qualifiers
FEATURES

```

source      1. .2385
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="II"
1. .2385
/gene="BLH1"
/locus_tag="AT2G26170"
/notes="homeodomain protein BLH1; transcription factor;
capable of heterodimerization with KNOT proteins; similar
to predicted homeodomain protein presented in GenBank
Accession Number AAD21463"
/codon_start=1
/product="BLH1-like homeodomain 1"
/protein_id="AAK43836.1"
/db_xref="GI:1387517"
/translation="MAAYFHNPPPEISAGSDGLQTLILNPTTYVQYTOQDNDNN
NNNNNNNNNNNNNNSSVFLDSHAFOPNASQFVGLPSGHEAAATADNIS
VLHGYPVRYSLYSGHVDPTQACACETPRAQGLSLTSSQQQQQHQHQPI
HVGFGSGHEDIRVSGSTGVTNGIANIVSKYLAQQLLDVYVADSDMNAKS
OLFSSKSGCNDKPVGESSAGAGGGAAGAKRPVELGTAEORIEOMKAKLS
NMLHEVQRVRYHQOMQVYSSFEQAGIGSAKSYSLAKTISQPRCLKEAJAQ
IKANKSLGDESVSGVREGRSLKFDVHLRQORALQQLGHIQHSNNARPRGL
PERAVSLRAWLFEHFLHPYKDSDKMLAKQTLGSLQVSNMFINARVRLKPMVEE
MYBEMKEQAKNMSMEKTELDQSDNEDSASKTSNOEKSPMADTYHNNHNDGLG
VTMGQCPKRLTSDETMQPFINADFSSEKLTWKILEERQGISDGYPFMGNFQY
QMDMSRFDVVSQDELMAORYSGNNVSLTLGLPHCDLSLSDHQQFMOTHRGIPG
RRVKGITEEYGPATNGSGSTTAAHSSAAAAAYNGMNIQNKRYVAQLLPDEVA"

ORIGIN
Query Match      11.9%; Score 324.6; DB 8; Length 2385;
Best Local Similarity 65.2%; Pred. No. 2.5e-55;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

Qy 758 TTTTAGGCTTAAGTATCTGAAGCTGCACAGAGCTTCTTGATGAAGTTGTAATTTG 817
Db 747 TTGTAGTCTCAAGTACTTGAAGGCACACAGAGCTTCTTGACGAAGTAGTCAACGCTG 806
Qy 818 TTGAAAAAGCATCAAGAGGAGATGATCAAAAGAGGATAATTCAATCAATAAAGATCAA 877
Db 807 ATTCCGATGACATGAACGCTAAATCCCACTATTCTCATCGAAAAGGGTAGTTGCGGAA 866
Qy 878 TGCCTTTGGCTAGTATGTCAACAATAATAGTTCTGCTGTGGTGTGAAGTAGCAGCAGGC 937
Db 867 ATGATAAACCTGTCGAGAGATCATCGCGCGCTGAGGAGAGAGTTCTCCGTTGGCGGAG 926
Qy 938 AGAAAAATG-----AGTTGCTGTTGAGCTTACACTGCTCAAGACAGAACTTCAAA 991
Db 927 CAGAAGCAGCGCGGAAACGTCGCTGGAGCTAGCACCGGAGAGAGACAAGAAATACAGA 986
Qy 992 TGAAGAAAGCCAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAGGTACAGACAGTACC 1051
Db 987 TGAAGAAAGCAAACTTAGTAAACATGCTTCATGAGGTGGAGAGATATAGACAGTACC 1046
Qy 1052 ATCAACAAATGCAAAATATGTTATATCATTTGAGCAAGTAGCAGGAAATGGATCAGCCA 1111
Db 1047 ACCAGCAGATGCAGATGCTGATCTCTTCTGTCGAGCAAGCGGAGGATAGGATCAGCGA 1106
Qy 1112 AATCATACATCAATTTAGCTTTGTCATCAATTTTGAAGCAATTCAGATGCTTAAGATG 1171
Db 1107 AGTCATACACGTCGCTAGCATTTGAACCCATATCAACACAGTTCCGTTGCTTGAAGAGG 1166
Qy 1172 CAATTGCTGAGCAGTAAAGCGCAGCAGCAAGAGTTAGTGTGAAGAGCAAGGCTTG--- 1227
Db 1167 CGATCGTGTGATCAATAAGCGGCCAACAGAGTCTTGGGAGGAGATTCAGTGTCTG 1226
Qy 1228 --GGAGGAAATCGAAGGCTCAAGACTCAAAATTTGGAACCATCATCTAAGGCAACAC 1285
Db 1227 GTGTTGGAGGTTTGAAGGGTTCGAGGCTCAAGTTCTGTTGGACCACTTGAAGCAGCAA 1286
Qy 1286 GCGGCGTGCACAGATAGGAATGATGCAAC-----CAATGCTTTGGAGACCCCAA 1336

```

```

Db 1287 GAGCTCTTCAACAACTGGGAATGATTAACATCTTCCAAATAATGCTTGAGAGCTCAAC 1346
Qy 1337 GAGGTTTACCTGAAAGAGCTGTCTCTGCTCCTGCTTGGCTTTTTCGAGCATTTTCTTC 1396
Db 1347 GTGGTCTCCAGAACGAGCGCTCTCAGTCTCTGCTTGGCTTCTTGAACACTTCTTC 1406
Qy 1397 ATCTTTACCCCAAGGATTCAGACAAATATGCTTGTGAAGCAAAACGGGCTTAACAGGA 1456
Db 1407 ATCCATACCCCTAAGGATTCGACAAAGCACATGCTAGCTAAGCAAAACAGGACTCACTGTA 1466
Qy 1457 GCAGGCTGCTTACTGCTTCAATAAATGCTGCTGATGCTGATGATGAGCCCAATGCTAGAG 1516
Db 1467 GCAGGCTGCTGAACTGCTTATTAACCGGAGAGTTCGTTATGGAACCAATGCTGGAGG 1526
Qy 1517 AAATGACTTTGGAAGAGTGAAGAATCAAGAAACAAACA 1555
Db 1527 AGATGTACATGAGGAATAATGAAGGACGAGCAAGAACA 1565

```

```

RESULT 10
AY072175
LOCUS      2572 bp      mRNA      linear      PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana putative homeodomain transcription factor
            (At2G35940) mRNA, complete cds.
ACCESSION  AY072175
VERSION     AY072175.1
KEYWORDS    FLI CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

```

```

REFERENCE  1 (bases 1 to 2572)
            Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
            Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
            Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
            Arabidopsis Full Length cDNA Clones
            Unpublished
            2 (bases 1 to 2572)
            Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
            Direct Submission

```

```

TITLE      Submitted (02-JAN-2002) Plant Gene Expression Center, 800 Buchanan
JOURNAL    Street, Albany, CA 94710, USA
AUTHORS    RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
            Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

```

The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGECC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGECC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source

1. .2572

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="2"

/clone="RAFL09-25-N04 (R18537)"

/ecotype="Columbia"

/notes="This clone is in a modified pBluescript vector (FLC-1) as a BamHI/XhoI insert."

1. .2572

/gene="At2g35940"

1. .205

/genes="At2g35940"

misc_difference 2

/genes="At2g35940"

/notes="compared to genomic sequence"

/replaces="c"

206. .2248

/gene="At2g35940"

/codon_start=1

/evidence=experimental

/product="putative homeodomain transcription factor"

/protein_id="AAI59997.1"

/db_xref="GI:18176173"

/translations="MAAYFHGNPPPEISAGSDGLQTLILMPTTVYVQYTCQNDNSNN NNSNNNTNTNTNNNSVFLDSHAPQPNASQOFVIGPLSGHEAASITAAADNIS VLGHPYPRVQYSLYGHVDPTFHQAACETPRAQQGLSLTSSQQQQQQHQQHQPI HVGFSGHGEDIIRVGSSTGSGVTNGIANLVSKYLKAAQELLDDEVNADSDMNAKS OLFPSKKSGCNDKPVGESAGAGSGGGAAGKRPVELGTAEBOELMKKAKLS NMLHEVEQRYQVHQOMQWISSFEQAAGIGSAKSYSLAKTISRQFRCLKEAIAGO IKAANKSGEEDSVSGVGRPEGRSLRFLPHLQRLQALQQLGMIQHPNSNNRPPQRL PERASVLRWLFEHLFHPYPKDSKHLMAKOTGLTRSQVSNWFNINARVLKMPVVEE MYMEENKEAKNMGSKETPLDQSDNEDSAS KSTSNQEKSPMADTNTYHNPNNHGDLEG VTGMCQSPKRLRTSGTQPIINADFSNEKLTMKLIERQIGIRSDGYPFMGNFQOY QMDMSRFDVSDQELMAQORYSGNNNGVSLTLGLPHCDLSLSTHOGFMQTHGIPIG RRVKIGETBEYGATINGSSTTTAHSAAAAAAYNGMNIQNKRYVAQLLPDFA"

2249. .2572

/gene="At2g35940"

misc_difference 2417. .2418

/genes="At2g35940"

/notes="compared to genomic sequence"

/replace="ttc"

misc_difference 2556

/genes="At2g35940"

/notes="not present in genomic sequence"

ORIGIN

Query Match 11.9%; Score 324.6; DB 8; Length 2572;

Best Local Similarity 65.2%; Pred. No. 2.5e-55;

Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

758 TTTTAGCTCTAAGTATCTGAAAGCTGCACAAAGAGCTTCTTGATGAGTTGTTAATATTG 817

774 TTGTAGCTCCAAGTACTTGAAGCAGCACAAGAGCTTCTTGACGAGTAGTCAACGCTG 833

818 TTGGAAGAACATCAAGGAGATGATCAAAAGAGGATAATCAATGAATAAAGATCAA 877

834 ATTCCGATGATCAAGCGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGA 893

878 TGCCTTTGGCTAGTGTGTCAACACTAATAGTTCTGCTGCTGCTGAAAGTAGCAGCAGGC 937

894 ATGATAAACCTGTTCGGGAATCATCGGCCGCGCTGGAGGAGAGGTTCGGTGGCGGAG 953

938 AGAAAAATG-----AAGTTGCTGTTGAGCTTACAACTGCTCAAGACAGAACTTCAA 991

954 CAGAGCAGCGCGGAAACGTCGCTGGAGCTAGCAGCGCAGAGAGACAGAAATACAGA 1013

992 TGAAGAAAGCAAGCTTCTTGCATGCTTGAAGAGGTGGAGCAAGAGTACAGACAGTACC 1051

1014 TGAAGAAAGCAAACTTAGTAACTAGTCTTATGAGGTGGAGCAGAGATATAGACAGTACC 1073

QY 1052 ATCACCAAAATGCAAAATAATTGTTATTATCATTTTGAGCAAGTAGCAGGAATTTGGATCAGCCA 1111

Db 1074 ACCAGCAGATGAGATGGTGTATCTCTTCGTTTGAGCAAGCGGAGGATAGGATCAGCGA 1133

QY 1112 AATCATACACTCAATTTAGCTTTTCATGCAATTTTGAAGCAATTTAGATGCTTAAAGGATG 1171

Db 1134 AGTCATACACGTCGCTAGCATTGAAACCATATCAAGACAGTTTCGCTTGCCTGAAAGAGG 1193

QY 1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGATTAGGTGAAGAGGAGGCTTG---- 1227

Db 1194 CGATCGCTGCTCAGATAAAAAGCGGCCCAACAAGAGTCTTGGGGAGGAGATTCAGTGTCTG 1253

QY 1228 --GGAGGGAAATCGAAGGCTCAAGACTCAAAATTTGTTGACCATCATCTAAAGCAACAAC 1285

Db 1254 GTGTTGGGAGGTTTGGAGGGTTCAGGCTCAAGTTCTGTGACCACTTGTGAGACAGCAAA 1313

QY 1286 GCGCGCTGCAACAGATAGGAATGATGCAAC-----CAAAATGCTTGGAGACCCCAAA 1336

Db 1314 GAGCTCTTCAACAACTGGGAATGATTCACATCTTCCAAATATGCTTGGAGACCTCAAC 1373

QY 1337 GAGTTTACTGAAAGAGCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1396

Db 1374 GTGCTCTCCGAAACGAGCGCTCTCAGTTCTCTCGCTCTTGGCTCTTTCGAACACTTTCTTC 1433

QY 1397 ATCCTTACCCAAAGGATTCAGACAAATCATGCTTCTAAGCAAAACGGGCTAAACAAGGA 1456

Db 1434 ATCCATACCTTAAGGATTCGGAACGACATGCTAGTAAAGCAAGACTCACTCTCGTA 1493

QY 1457 GCCAGGTGCTAACTGGTTTCTAATAATGCTCGAGTTTCGATTTATGAAAGCCAAATGGTAGAAG 1516

Db 1494 GCCAGGTGCGAATCGTTTATAAAACGCGAGAGTTTCGGTTATGGAACCAATGGTGGAGG 1553

QY 1517 AATGTACTTGGAGAGTGAAGATCAAGAACAAAACA 1555

Db 1554 AGATGTACATGGAGGAATGAAGGAGCAGCGCAAGAACA 1592

RESULT 11

AY099854

LOCUS

Arabidopsis thaliana putative homeodomain transcription factor (At2g35940) mRNA, complete cds.

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 2578)

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Bath,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayaishizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.

Direct Submission

Submitted (24-APR-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayaishizaki,Y. and Shinozaki,K.

The Salk, Stanford, RGC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,

Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K., Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologos, A. and Davis, R. W.

Nguyen, M. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinoraki, K. (RIKEN GSC) and Davis, R. W. (SSP/Stanford) contributed equally to this work as PIs.

Location/Qualifiers

1..2578

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="2"

/clone="RAPL09-94-K12 (R25364)"

/ecotype="Columbia"

/note="This clone is in pBluescript"

1..2578

/gene="At2g35940"

/note="synonym: F11F19.15"

198..2240

/gene="At2g35940"

/codon_start=1

/product="putative homeodomain transcription factor"

/protein_id="AA020705.1"

/db_xref="GI:20466776"

/translation="MAAYFGNPEISAGSDGGIQLTILMNPITYVQYTQDDNSNNNSNNNSNNNTNTNNNSFVLDHAPQNPASQOVGVPLSGHEAASITAADNISVLHGYPVVOYSLYGHSDVPTHQOACGTAAQOGLSLTSSQOQQOQHOHQHPIHVGFGSHGDIKRVGSGTGVNTGCIANLVSKYLKAQELLDVNVNADSDMNAKQLPSSKGCSCNDKPVGESAGAGGSGGGAAGKRPVELGTAEKQELQMKAKULMLHUEBQRVRYHQWQWIVISFPEQAAGISGSAKSYTSLAKTISRQFCLKEAIAQGLIKAANKSLGBEDSVQGVREGSLKFDVHLLQORALQQLGMIQHPSSNNAWPRQGLPEBAAVSLRAFLFEHLPHYPLQDSKDHLAKTGLTRSOVSNWFINARVLKMPWEE MYMEEMKQAKNMGSEKTPLDOSNDSASKTSNQBSPMADNTNHNPNHNGDLEGVTMQGSPKLRYSDETWMOPINADPSSNEKLPMKLLIEEPOGIRSGGYPPMGNGPOYQTMGSRFVDSQBELMAQRYSGNNNGVSLTGLPLCHDSLSSHHTHGHGPIPIGRVKMISGETEYGTAGTNGSGSSTTTTAHSSAAAAAYNGMNIONKRYVAQLLPDFVA

ORIGIN

Query Match	Best Local Similarity	Score	324.6;	DB 8;	Length	2578;
Matches	534;	Conservative	0;	Mismatches	284;	Indels
Qy	758	TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTTAATTG	817			
Db	766	TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG	825			
Qy	818	TTGAAAAAGCATCAAGAGAGATGATCAAAAGAAGGATATTCATGAATAAAGAATCAA	877			
Db	826	ATTCGCGATGATCAAGCGCTAAATCCCAACTATTCTCATGAAAAAGGAGTAGTTCGGGA	885			
Qy	878	TGCTTTTGGCTAGTGATGTCAACACTAATAGTCTCTGTTGGTGGTGAAGTAGCAGCAGGC	937			
Db	886	ATGATAAACCTGTCCGGAATATCATCGCGCGCGCTGGAGAGAAAGGTTCCGGTGGCGGAG	945			
Qy	938	AGAAAAATG-----AAGTTGCTCTTGAGCTTACAACTGTCTCAAGACAAGAACTTCAAA	991			
Db	946	CAGAAGCAGCCGGAAACGTCGGTGGAGCTAGGCACGGCAGAGACAGAATAACAGA	1005			
Qy	992	TGAAAAAGCCAGCTTCTTGCCCATGTTTGAAGAGGTGGAGCAAGGTTACAGACAGTACC	1051			
Db	1006	TGAAGAAGCAAACTTAGTAAACATGCTTCATGAGTGGAGCAGAGATATAGACAGTACC	1065			
Qy	1052	ATCACAAATGCAATAATTGTTATTATTCATTGGAGCAAGTAGCAGGAATTCGATCAGCCA	1111			
Db	1066	ACCAGCAGATGCAGATGGTATCTCTTCGTTGAGCAAGCGCAGGATAGGATCAGCGA	1125			
Qy	1112	AATCATACACTCAATTAGCTTTGATGTCAAATTCGAAAGCAATTCAGATGCTCAAGGATG	1171			
Db	1126	AGTCATACAGTCGCTAGCATTTGAAACCATATCAAGACAGTTCCTGTTGCTTGAAGAGG	1185			
Qy	1172	CAATTGCTGAGCAAGTAAAGCGCAGCAGAGAGTTTTAGTGTGAAGAGGAAGGCTTG----	1227			

Db	1186	CGATCGCTGCTGAGATAAAAGCGGCCAAACAAGAGTCTTTGGGGGAGGAAGATTCTAGTGTCTG	1241
Qy	1228	--GGAGGGGAAAATCGAAGGCTCAAGACTCAAAATTTTGTGGACCAATCATCTTAAGGCAACAAC	1285
Db	1246	GTGTTGGAGGTTTGGGGGTCGAGGCTCAAGTTCTGTGGACCAACCACCTTGAGACAGCAA	1305
Qy	1286	GCGGCTGCACAGATAGGAATGATGCAAC-----CAATGCTTGGAGACCCCAAA	1336
Db	1306	GAGCTCTTCAACAACATGGGAATGATTCAACATCTTCCATAATAGTCTTGAGACCTCAAC	1365
Qy	1337	GAGGTTTACCTGTAAGAGCTGCTCTGTCCTTCGTCTGCTTGGCTTTTCGAGCATTTTCTTC	1396
Db	1366	GTGGTCTCCAGAACGAGCGCTCTCAGTTCTCCGTCTTGGCTCTTCGAAACATTTTCTTC	1425
Qy	1397	ATCCTTTACCCAAAGGATTTCAGACAAAATCATGTTCTTAAGCAAAACGGGCTAACAGGA	1456
Db	1426	ATCCATACCCTTAGGATTTCGGAACGACACATGCTAGCTAAGCAACAGGACTCACTCGTA	1485
Qy	1457	GCCAGGTGTTAATCTGGTTTCATTAATGCTCGAGTTTCGATTATGGAAGCCAAATGGTAGAG	1516
Db	1486	GCCAGGTGTCGAACCTGGTTTATTAACCGGAGAGTTCGGTTATGGAACCAACATGGTGGAGG	1545
Qy	1517	AAATGTACTTGGGAAGTGAAGTAAGTAACGTAACGACAAACA	1555
Db	1546	AGATGTACATGGGAGGAATGAAGGAGCGGCAAGAACA	1584
RESULT 12			
AK104872			
LOCUS	AK104872	1896 bp	linear
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:001-044-C01, full insert sequence.		
ACCESSION	AK104872		
VERSION	AK104872.1	GI:32990081	
KEYWORDS	FLU_CDNA; oligo capping.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
REFERENCE	1		
AUTHORS	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team., Kikuchi,S., Satoh,K., Negata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., Riken:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
JOURNAL	Science 301 (5631), 376-379 (2003)		
MEDLINE	22752273		
PUBMED	12869764		
REFERENCE	2 (bases 1 to 1896)		
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hizamoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,		

Masuda, H., Matubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Oota, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-338-7007, Fax: 81-29-338-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oota, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Yahashizaki, Y.

Location/Qualifiers
1. .1896
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xrefs="taxon:39947"
/clone="001-044-C01"

FEATURES

source
Query Match 10.8%; Score 295.6; DB 8; Length 1896;
Best Local Similarity 60.3%; Pred. No. 1.9e-49;
Matches 528; Conservative 0; Mismatches 339; Indels 9; Gaps 2;

ORIGIN

765 CTCTAAGTATCTGAAGCTGCACAAGAGCTTCTTGATGAAGTGTGTTAATATTTGTGAAA 824
|||
301 CTCAAGGTATCTCAAGSCAGCACAAGAAGCTGCTTGATGAGTGTGTTGGAAGAG 360
|||
825 AAGCATCAAGGAGATGATCAAGAAGAGGTAATTCATCAATAAGAATCAATGCGCTTT 884
|||
361 CATATAACAAGAGCTCAGAAAGAGAGGTTGTAATCAGGAAAGAGCATGGCAAA----- 415
|||
885 GGCTAGTGATGTCAACACTAATAGTCTCGTGGTGGTGAAGTAGCAGCAGGCAAGAAA 944
|||
416 -GAGACTGATGGAGGCGCAAAAGTGGGGTGTATCTTCTTAATCCACAAGAATCTGGCG 474
|||
945 TGAAGTTGCTGTGAGCTTACAACCTGCTCAAGACAGAACTTCAATGAAGAAAGCCAA 1004
|||

Db 475 TAATGAGCAGCACTGAGCTTTTCCACTGCTGAGAAGCAAGAGCTTTCAGACACAAATATGCAAA 534
Qy 1005 GCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAAGGTATACAGACAGTACCATCACCATAATGCA 1064
Db 535 ACTGATGGCAATGTTGGATGAGGTGGACCGGAATAACAAGCATATTATCCACCAATGCA 594
Qy 1065 AATAATTGTATTATCATTTTGGAGCAAGTACAGCAATTTGGATCAGCCCAAAATCATACACTCA 1124
Db 595 AACCGTGGTTTTCATCGTTTGATGTGGTAGTCGACCAAGGATCTGCGAAGCCTTACACAGC 654
Qy 1125 ATTAGCTTTGTCATGCAATTTGGAAGCAATTCAGATGCTTAAAGATGCAATTTGCTGAGCA 1184
Db 655 AGTTGCTCTTCAGACAAATCTCAGGCACCTTCGGTGGCTGAAGGATGCTTATCAATGATCA 714
Qy 1185 AGTAAGGCGCAGCAGCAAGAGTTTAGTGAAGNAGGCTTTGGGAGGGAATAATCGAAGG 1244
Db 715 GATCAATGTTATCAGGAAAAAGCTTGGCGAGGAAGAACTCATCTGGCAAGGAGGGA 774
Qy 1245 CTCAA---GACTCAAAATTTGTGGACCATCATCTTAAGGCAACACGCGCTGCAACAGAT 1301
Db 775 ATTAACCCGCTCCGGTACATTGATCAGCAGCTAAGGCAACACGCTTTCACACAGTA 834
Qy 1302 AGGAATGATGCAACCAAAATGCTTGGAGACCCCAAGAGGTTTACCTGAAAGAGCTGTCTC 1361
Db 835 CGGTATGATTTCACAAAAAGCGCTGGAGACCAAGAGGAGCTGCTGAAAACTCTGTGTAC 894
Qy 1362 TGCTCTTCTGCTGGCTTTTCGAGCATTTTCTTCACTCTTACCCAAAGGATTCACACAA 1421
Db 895 AATTCTTCGCGCTTGGCTGTTTGAACACATTCCTCCATCCGTAACCCAAAAGATTCGAGAA 954
Qy 1422 AATCATGCTTGTGAAGCAACGCGGCTAAACAAGGAGCAGGCTGTCTTAACTGGTTCATAA 1481
Db 955 GTTAATGCTTGTAGCAGACTGGCTGACTAGGAGTCAGATTTTCAATTTGGTTCATAA 1014
Qy 1482 TGCTGAGTTGCAATTAAGAACCAATGTTAGAGAAATGTTACTTGGAGAGAGTGAAGAA 1541
Db 1015 TGCCCGTGTCCGCTTTTGGAAACCAATGATGAAGACATGTCAAAAGAGAGATTGGTGA 1074
Qy 1542 TCAAGAACAAACAGTACTTAATCTTCAGAGATTAACAAACAAAGAGACCAATATAG 1601
Db 1075 TTTAGAGCAAGACTTCCAACTCTTCTGACATGCAACCAAGAGTAAGATTAATATAGC 1134
Qy 1602 TGCTCCAAATGAAGAAACATCCCAATTTACTAG 1637
Db 1135 ATCTTCAGAGACAGAGGAGTCTCAAAATCCAG 1170

RESULT 13

AK121356
LOCUS
DEFINITION
insert sequence.
AK121356
VERSION
FLI_CDNA; CAP trapper.
KEYWORDS
Oryza sativa (japonica cultivar-group)
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1
The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team., Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kuroseki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;

Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Aizawa,T., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Oeato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hirooka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Inamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakahama,Y., Nakamura,M., Naniki,T., Nariawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Ootomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice
Unpublished
3 (bases 1 to 2476)
Kikuchi,S.
Direct Submission
Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp)
Tel:81-29-838-7007, Fax:81-29-838-7007
This clone is one of the 32K full-length cDNA clones from japonica rice.
URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NTAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Yamamoto,M. and Nakahama,Y.
FAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Iehibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Nariawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.
Location/Qualifiers

FEATURES

source
1. 2476
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023122008"

ORIGIN

Query Match 10.8%; Score 295.6; DB 8; Length 2476;
Best Local Similarity 60.3%; Pred. No. 1.9e-49;
Matches 528; Conservative 0; Mismatches 339; Indels 9; Gaps 2;
Qy 765 CTCTAAGTATCTGAAGCTGCACAGAGCTTCTTTGATGAAGTTGTTAATATTTGTTGAAA 824
Db 1339 CTCAAGGTATCTCAAGCAGCAGCAAGAACTGCTTTGATGAGTCTGTTGGAAGAG 1398
Qy 825 AAGCATCAAGGAGATGATCAAAAGAGAGTAATTTCAATGAATAAAGAAATCAATGCCTTT 884
Db 1399 CATAAAAACAAAGGCTCAGAAAGAGAAAGTTGAATCAGGAAAGCAGATGSCAAA----- 1453
Qy 885 GGCTAGTGTGTCAACACTAATAGTTCTGCTGGTGGTGAAGTAGCAGCAGGCGAGAAAAA 944
Db 1454 -GAGACTGATGGAGGGCCAAAAAGTGGGGTGTATCTTCTAATCCACAAGATCTGGCG 1512
Qy 945 TGAAGTTGCTGTGTAGCTTCAACTGCTCAAGACAAAGAACTTCAAAATGAAAAAGCCAA 1004
Db 1513 TAATGACGACCTGAGCTTTCCACTGCTGAGAAAGCAAGAGCTTCAGAAACAAATGGCAAA 1572
Qy 1005 GCTTCTTGCCATGCTTGAAGAGGTGGAGCAAGGTACAGACAGTACCATCACCAATGCA 1064
Db 1573 ACTGATGGCAATGTTGGATGAGGTGGACCGGAATACAAAGCATTTATTCACCAAAATGCA 1632
Qy 1065 AATAATTGTATTATCATTTTCAGCAAGTAGCAGGAATTTGGATCAGCCAAATCATACACTCA 1124
Db 1633 AACGTGTTTCATCGTTTGATGTGTAGTGTGACCAAGGATCTGCGAAGGCTTACACAGC 1692
Qy 1125 ATTAGCTTTGCATGCAATTTGGAAGCAATTCAGATGCTTAAAGGATCAATTTGCTGAGCA 1184
Db 1693 AGTTGCTTTTCAGACAACTCTCAGGCATTTTCGGTGCCTGAAGGATCTTCAATGATCA 1752
Qy 1185 AGTAAGGCGACGACGACAGAGTTTAGTGAAGGAAGGCTTTGGGAGGGAATTCGAAGG 1244
Db 1753 GATCAATGTTTATCAGGAAAAAGCTTGGCGAGGAAGAACTCATCTGGCAAGGAGGAAA 1812
Qy 1245 CTCAA---GACTCAAAATTTGTGACCAATCATCTAAGGCAACAAACGCGCGCTGCAACAGAT 1301
Db 1813 ATTAAACCGCTCGGTACATTGATCAGCAGCTTAAGCAACACAGCTGTTTCCACAGTA 1872
Qy 1302 AGGAATGATCAACCAATGCTTTGGAGACCCCAAGAGGTTTACCTGAAAAGAGCTGTCTC 1361
Db 1873 CGGTATGATTTCCACAAAAACGCTGGAGACCACAGAGGGGACTGCCTGAAAACCTCTGTTAC 1932
Qy 1362 TGTCCTCGTGTGCTTTCGAGCAATTTTCATCCTTACCCAAAGGATTCAGACAA 1421
Db 1933 AATTCTCGGCGCTTGGCTGTTTGAACACTTCTCCCTCCATCCGTAACCAAGATTCGAGAA 1992
Qy 1422 AATCATGCTTGTCTAAGCAACACGGGGCTTAAACAGGAGCGAGTGTCTAACTGGTTTCAATA 1481
Db 1993 GTTAATGCTTGTCTAGGACAGCTGGCTGACTAGAGTCAGATTCGAAATTTGTTTCATAA 2052
Qy 1482 TGCTCGAGTTTCGATTTAGGAAGCCAATTTGGTAGAAGAAATGATCTTTGGAAGAGTGAAGAA 1541
Db 2053 TGCCCGTGTCCGCTTTTGGAAACCAATGATTGAAGACATGTACAAGAAGAGATTTGGTGA 2112
Qy 1542 TCAAGAACAAAACAGTACTTAATCTTCAGAGAGATTAACAAAACAAAGAGACCAATATAG 1601
Db 2113 TTTTAGAGCAAGACTTCCAACTCTTCTCTGCAATTCGACCAGGAAGTAAGGATAAAATGGC 2172
Qy 1602 TGCTCCAAATGAAGAGAGAAACATCAATTTATTACTAG 1637
Db 2173 ATCTTCAGAGACAGAGGATCTCAAAAACACTCCAG 2208

RESULT 14	AK070465	3030 bp	mRNA	linear	PLN 24-JUL-2003
LOCUS	AK070465				
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J023054P21, full insert sequence.				
ACCESSION	AK070465				
VERSION	AK070465.1	GI:32980489			
KEYWORDS	FLI CDNA; CAP trapper.				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Eurykaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Eriarthraloideae; Oryzaceae; Oryza.				
REFERENCE	1				
AUTHORS	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Nariikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.				
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice				
JOURNAL	Science 301 (5631), 376-379 (2003)				
MEDLINE	22752273				
PUBMED	12869764				
REFERENCE	2 (bases 1 to 3030)				
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanaeawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Nariikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otono,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shingawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-DEC-2001) Shohji Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)				
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice. URL : http://cdna01.dna.affrc.go.jp/cdna/ NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M. FAIS Genome Sequencing & Analysis Group: Otono,Y., Iida,Y.,				
FEATURES	source				
ORIGIN	1..3030				
Query Match	10.8%; Score 295.6; DB 8; Length 3030;				
Best Local Similarity	60.3%; Pred. No. 1.9e-49;				
Mismatches	528; Conservative 0; Mismatches 339; Indels 9; Gaps 2;				
765	CTCTAGTATCTGAAGCTGCACAGAGCTTCTTGATGAAGTTGTAAATATTTGTTGAAA 824				
1394	CTCAAGGTATCTCAAGGGCAGCAGAACTGCTTGATAGGTCGTGAGTGTGGAAGAG 1453				
825	AAGCATCAAGGAGATGATCAAAAGAGGATTAATTCATGAATAAAGATCAATGCTTT 894				
1454	CATAAACAAAGGCTCAGAAAGAGAGGTTGAATCAGGAAAGCAGATGGCAAA---- 1508				
885	GGCTAGTGTATCAACACTAATAGTTCTGTGTGTGTGAAGTAGCAGCAGGCAAGAAA 944				
1509	-GAGACTGATGAGGGCCAAAAGTGGAGGTGTATCTTCTAATCCACAAAGATCTGGCGC 1567				
945	TGAAGTTGCTGTGTAGCTTACAACTGCTCAAGACAGAACTTCAATGAAAGAACCCAA 1004				
1568	TAATGAGCACCCTGAGCTTTCACCTGCTGAGAGCAAGAGCTTCAGACAAATGCGCAA 1627				
1005	GCTTCTTGCCATGCTTGAAGAGGTGGAGCAAGGTACAGACAGTACCATCCAAATGCA 1064				
1628	ACTGATGGCAATGTTGGATGAGGTGGACCGGAAATACAGCATTTATACCACCAATGCA 1687				
1065	ATAATGTTATATCAATTTGAGCAATAGCAGGAATTTGATCAGCCAAATCATACACTCA 1124				
1688	AACCGTGGTTTCATCGTTTGTGTAGTGTGAGCAGGATCTGCGAAGCCTTTACACAGC 1747				
1125	ATTAGCTTTGTCATGCAATTTCCGAAGCAATTCAGATGCTTAAAGGATGCAATTCCTGAGCA 1184				
1748	AGTTGCTCTTCAGACAACTTCACGGCCTTTGCGTGCCTTGAAGGATGCTATCAATGATCA 1807				
1185	AGTAAAGGCGCAGCAGCAAGAGTTTATAGTGAAGAGGAGGCTTTGGGAGGGAATTCGAAGG 1244				
1808	GATCAATGTTATCAGGAAAAAGCTTGGCGAGGAAGAGAACTCATCTGGCAAGGAGGAAA 1867				
1245	CTCAA---GACTCAAAATTTGTGACCATCATCTAAGGCAACACGCGCCTGCAACAGAT 1301				
1868	ATTAAACCCGCTCCGGTACATTTGATCAGCAGCTTAAGGCAACACGCTTTTCCACAGTA 1927				
1302	AGGAATGATGCAACCAAAATGCTTTGGAGAGCCCAAGAGAGGTTTACCTGGAAGAGCTGCTC 1361				
1928	CGTATGATTTCCAAACAAAGCCCTGGAGACACAGAGGGAAGCTGCTTGAAGAACTCTGTAC 1987				
1362	TGTCCTTCTGCTGGCTTTTTCGAGCAATTTTCTTCATCTTACCCCAAGGATTCAGACAA 1421				

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 05:17:43 ; Search time 1407 Seconds
(without alignments)
11507.102 Million cell updates/sec

Title: US-10-624-201a-1
Perfect score: 2735
Sequence: 1 catcgagagataaaatata.....gcaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	324.6	11.9	2043	6	ABZ13532 Arabidops
2	324.6	11.9	2043	9	ABZ23149 Arabidops
3	324.6	11.9	2385	10	ADD30303 Plant yie
4	324.6	11.9	2385	12	AD144208 Plant tra
5	278.8	10.2	1929	12	ADQ37104 Cell prol
6	261	9.5	2031	3	AAC36460 Arabidops
7	259.4	9.5	1449	3	AAC43406 Arabidops
8	259.4	9.5	1449	6	ADG88236 A. thalia
9	244.4	8.9	1971	12	ADO62706 Transcrip
10	237.6	8.7	706	6	ABK82114 DNA encod
11	234.6	8.6	2025	3	AAC39138 Arabidops
12	234.4	8.6	1575	4	AD06493 Arabidops
13	234.4	8.6	1575	12	ADO61968 Transcrip
14	231.8	8.5	1824	3	AAC36745 Arabidops
15	231.8	8.5	2458	12	ADO62707 Transcrip
16	230.2	8.4	1617	12	ADO63042 Transcrip
17	225.4	8.2	1599	6	ABZ14349 Arabidops
18	224	8.2	1884	6	ABZ13100 Arabidops
19	224	8.2	1983	4	AD06494 Arabidops
20	224	8.2	1983	10	ADC46618 Thalecres

21	224	8.2	1983	10	ADD30786	Add30786 Plant yie
22	224	8.2	1983	12	AD144018	Adi44018 Plant tra
23	224	8.2	1983	12	ADO02224	Ado02224 Thalecres
24	222.6	8.1	2545	4	AAD06479	Aad06479 Arabidops
25	222.6	8.1	2545	10	AD837296	Ad837296 Plant yie
26	222.6	8.1	2545	12	AD143866	Adi43866 Plant tra
27	222.6	8.1	2545	12	ADO01854	Ado01854 Thalecres
28	213.4	7.8	1485	3	AAC50215	Aac50215 Arabidops
29	207	7.6	1296	12	ADO61794	Ado61794 Transcrip
30	201.4	7.4	626	5	AAH87740	Aah87740 Peppermin
31	187.2	6.8	1836	6	ABZ12846	Abz12846 Arabidops
32	187.2	6.8	1836	6	ADG88038	Adg88038 A. thalia
33	187.2	6.8	2405	4	AAD06449	Aad06449 Arabidops
34	187.2	6.8	2405	10	ADE37116	Ade37116 Plant yie
35	187.2	6.8	2405	12	AD143868	Adi43868 Plant tra
36	187.2	6.8	2405	12	ADO01856	Ado01856 Thalecres
37	183.8	6.7	529	3	AAC55946	Aac55946 Eucalyptu
38	182.8	6.7	662	13	ADR65754	Adr65754 Cotton cd
39	180.4	6.6	592	13	ACN59439	Acn59439 Cotton gy
40	176.8	6.5	2034	3	AAC50226	Aac50226 Arabidops
41	170.2	6.2	276	12	ADQ05068	Adq05068 Soybean h
42	166	6.1	894	12	ADQ37116	Adq37116 Cell prol
43	164.6	6.0	343	3	AAC56856	Aac56856 Pinus rad
44	159.8	5.8	323	3	AAC56287	Aac56287 Pinus rad
45	158	5.8	833	3	AAC34112	Aac34112 Arabidops

ALIGNMENTS

RESULT 1
ABZ13532
ID ABZ13532 standard; DNA; 2043 BP.

XX ABZ13532;
XX
XX 21-JAN-2003 (first entry)
XX
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1337.
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana.
XX
XX WO200216655-A2.
XX
XX 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US026685.
XX
XX 24-AUG-2000; 2000US-0227866P.
XX 26-JAN-2001; 2001US-0264647P.
XX 22-JUN-2001; 2001US-0300111P.
XX (SCRI) SCRIPPS RES INST.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
XX
XX WPI; 2002-304127/34.

Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

Claim 144; SEQ ID NO 1337; 577pp + Sequence Listing; English.

The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants

CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 2043 BP: 627 A; 507 C; 507 G; 402 T; 0 U; 0 Other;

AC	ADB23149;
XX	
XX	20-NOV-2003 (first entry)
XX	
DE	Environmental stress-responsive promoter-related DNA, SEQ ID 153.
XX	
XX	Plant; environmental stress; promoter; gene; ds.
KW	
XX	
XX	Arabidopsis thaliana.
OS	
XX	WO2003044190-A1.
PN	
XX	
PD	30-MAY-2003.
XX	
XX	15-NOV-2002; 2002WO-JP011955.
PF	
XX	
XX	19-NOV-2001; 2001JP-00353038.
PR	
XX	29-JAN-2002; 2002JP-00020329.
XX	
XX	
PA	(RIKE) RIKEN KK.
XX	
XX	Shinozaki K, Seki M, Fujita M;
PI	
XX	
XX	WPI; 2003-468771/44.
DR	
XX	P-PSDB; ADB23150.
XX	
XX	
PT	Environmental stress-responsive promoters, useful in constructing and
PT	breeding stress-resistant plants for applications in agriculture.
XX	
XX	Claim 10; Page 470-475; 520pp; Japanese.
XX	
XX	
CC	The present invention relates to novel environmental stress-responsive
CC	promoters (ADB23163-ADB23252) from Arabidopsis thaliana. The promoters
CC	are useful in constructing and breeding stress-resistant plants for
CC	applications in agriculture and horticulture. The present sequence was
CC	used to illustrate the invention.
XX	
XX	
SO	Sequence 2043 Bp; 627 A; 507 C; 507 G; 402 T; 0 U; 0 Other;

Query Match	11.9%	Score 324.6	DB 9	Length 2043
Best Local Similarity	65.2%	Pred. No. 3.8e-64		
Matches 534	Conservative 0	Mismatches 264	Indels 21	Gaps 3
Qy	758	TTTGTAGGCTCTAAGTATCTGAAAGCTGCACAGAGCTTCTTGATGAAGTGTGTTAATATTG	817	
Db	569	TTGTTAGCTCCAAGTACTTGAAGCGACACAGAGCTTCTTGACGAAGTAGTCAACGCTG	628	
Qy	818	TTTGAAAAAGCATCAAAGGAGATGATCAAAGAAGGATATTTCAATGAATAAAGAATCAA	877	
Db	629	ATTCCGATGACATGAACGCTAAATCCCACTATTCATCGAAAAAGGGTAGTTGCGGAA	688	
Qy	878	TGCTTTTGGCTAGTATGTCAACACTAATAGTTCTGGTGGTGTGTAAGTAGCAGCAGGC	937	
Db	689	ATGATAAACCTGTCGGAGAAATCATCGCCGCGCTCGAGGAGAGGTTCCCGTGGCGGAG	748	
Qy	938	AGAAAAATG-----AAGTTGCTGTTGAGCTTACAACCTGCTCAAAGACAAAGAACTTCAA	991	
Db	749	CAGAAGCAGCCGGAAACGTCGCGTGAGCTAGGCACGGCAGAGAGACAAGAATAACAGA	808	
Qy	992	TGAAAAAGCCAAGCTTCTTGCCCATGCTTGAAGAGGTGGAGCAAAAGTACAGACAGTACC	1051	
Db	809	TGAAGAAAGCAAACTTAGTAAACATGCTTCATGAGTGGAGCAGAGATATAGACAGTACC	868	
Qy	1052	ATCACAAATGCAAAATAATTTGTTATATCATTTGAGCAAGTAGCAGGAATTTGGATCAGCCA	1111	
Db	869	ACCAGCAGATGCAGATGGTGATCTCTCTGTTTCGAGCAAGCGCGAGGATAGGATCAGCGA	928	
Qy	1112	AATCATCACTCAATTAGCTTTTCGATGCAATTTTCGAAGCAATTCAGATGCTCTAAAGATG	1171	
Db	929	AGTCATACAGCTCGCTAGCAATGAAACCATAATCAAGACAGTTTCCTCGTTGCTTGAAGAGG	988	
Qy	1172	CAATTGCTGAGCAAGTAAAGCGCACGCAAGAGTTTATAGTGAAGAGGAAGGGTTG----	1227	

RESULT 2
ADB23149
ID ADB23149 standard; DNA; 2043 BP.
XX

Db 989 CGATCGTGTGTCAGATAAAGCGGCCCAACAGAGCTTCTGGGAGGAAGATTTCAGTGTCTG 1048
 Qy 1228 --GGAGGGAATAACGAAGGCTCAAGACTCAAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
 Db 1049 GTGTTGGAGGTTTGGAGGTCGAGGCTCAAGTTTCGTGGACCACTTGAGACAGCAAA 1108
 Qy 1286 GCGGCTGCAACAGATAGGATGATGCAAC-----CAATGCTTGGAGACCCCAAA 1336
 Db 1109 GAGCTCTTCAACAACACTGGGAATGATTCACATCTTCCCAATAATGCTTGGAGACCTCAAC 1168
 Qy 1337 GAGGTTTACCTGAAAGAGTGTCTGTCTCTTCTGCTTGTGCTTTTCGAGCATTTTCTTC 1396
 Db 1169 GTGCTCTCCAGACAGCGCTCTCAGTTCTCCGTGTTGGCTTTCGACACTTCTTC 1228
 Qy 1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGTCTAAGCAAAACGGGCTAACAGGA 1456
 Db 1229 ATCCATACCTTAAGGATTCGACAAAGCACATGCTAGCTAAGCAACAGGACTCACTCGTA 1288
 Qy 1457 GCCAGGTGTCTAATGTTTATATAATGCTCGAGTTTCGATTATGGAACCAATGGTAGAAG 1516
 Db 1289 GCCAGGTGTCAACTGGTTTATATAACGCGAGGTTTCGGTTATGGAACCAATGGTGAGG 1348
 Qy 1517 AATGTACTTGGGAAGAGTGAAGAAATCAAGAACAAACA 1555
 Db 1349 AGATGTACATGGAGAAATGAAGAGCAGGCAAGAACA 1387

RESULT 3

ADD30303
 ID ADD30303 standard; cDNA; 2385 BP.

AC ADD30303;

XX 15-JAN-2004 (first entry)

XX Plant yield-related polynucleotide clone G1589.

XX ds; transcription factor; transgenic plant; growth rate; senescence;
 KW seed germination rate; plant vigor; seedling vigor.

XX Arabidopsis thaliana.

XX WO2003013227-A2.

XX 20-FEB-2003.

XX 09-AUG-2002; 2002WO-US025805.

XX 09-AUG-2001; 2001US-0310847P.

XX 19-NOV-2001; 2001US-0336049P.

XX 11-DEC-2001; 2001US-0338692P.

XX 14-JUN-2002; 2002US-00171468.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;

XX Pilgrim M, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;

XX Broun PE;

XX WPI; 2003-248221/24.

XX P-PSDB; ADD30304.

XX New plant transcription factor polynucleotides and polypeptides, useful

XX in producing transgenic plants with commercially valuable properties,
 PT such as an alteration in a plant growth characteristic, e.g. growth rate
 PT or apomixis.

XX Disclosure; SEQ ID NO 332; 454pp; English.

XX The invention relates to a number of isolated Arabidopsis thaliana cDNA
 CC sequences and their encoded proteins which are especially transcription
 CC factor related cDNA's and proteins. The isolated or recombinant plant
 CC transcription factor polynucleotides and polypeptides are useful in

CC producing transgenic plants with commercially valuable properties, i.e.
 CC modified or altered desirable traits as compared to a reference plant,
 CC such as an alteration in a plant growth characteristic, e.g. growth rate,
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and
 CC flower senescence. Sequence information related to the polynucleotides
 CC and polypeptides can also be used in bioinformatic search methods. The
 CC transgenic plant is useful for growing a progeny plant from a parent
 CC plant. This sequence represents one of the cDNAs of the invention.

XX Sequence 2385 BP; 748 A; 578 C; 565 G; 494 T; 0 U; 0 Other;

Query Match 11.9%; Score 324.6; DB 10; Length 2385;

Best Local Similarity 65.2%; Pred. No. 4e-64;

Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

Qy 758 TTTTGGCTTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
 Db 747 TTGTTAGCTCAAGTACTTGAAGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACCTG 806
 Qy 818 TTGGAAGAAAGCATCAAGGAGATGATCAAAAGAGGATAATTCATGAATAAAGAAATCAA 877
 Db 807 ATTCCGATGACATGAACGCTAAATCCCACTATTCTCATGAAAAAGGGTAGTTGCGAA 866
 Qy 878 TGCCTTTGGCTAGTGTGTCACACTAATAGTTCTGCTGCTGTAAGTAGCAGCAGGC 937
 Db 867 ATGATAAACCTGTGCGAGAATCATCGCGCGCTGAGGAGAGGTTCCGGTGGCGGAG 926
 Qy 938 AGAAAAATG-----AAGTTGCTGTTGAGCTTCAACTGTCAAGAGCAAGAACTTCAAA 991
 Db 927 CAGAGCAGCGGGAACGTCCTGCTGAGCTAGCGCGGAGAGAGACAGAAATACAGA 986
 Qy 992 TGAAGAAAGCAGCTTCTTGCCATGCTTGAAGAGGTGAGAGCAAGGTACAGACAGTACC 1051
 Db 987 TGAAGAAAGCAAACTTAGTAACATGCTTCATGAGGTGAGAGAGATATAGACAGTACC 1046
 Qy 1052 ATCACCAGATGCAAAATAATTGTTATTATTCATTTCAGCAAGTAGCAGGAAATGGATCAGCCA 1111
 Db 1047 ACCAGAGATGAGATGGTGATCTCTTCGTCGAGCAAGCGGAGGATAGGATCAGCGA 1106
 Qy 1112 AATCATACACTCAATTAGCTTTGCATGCAATTTTCAAGCAATTCAGATGCTTAAGGATG 1171
 Db 1107 AGTCATACAGTCGCTAGCATTGAAACCATATCAAGACAGTTCCTGTTGTTGAAGAGG 1166
 Qy 1172 CAATTGCTGAGCAAGTAAGGCGAGCAAGAGTTTAGGTGAAGAGAGAGGCTTG---- 1227
 Db 1167 CGATCGCTGTCAGATAAAAGCGGCCCAACAAGAGTCTTGGGAGGAGAGATTTCAGTGTCTG 1236
 Qy 1228 --GGAGGAAATCGAAGGCTCAAGACTCAAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
 Db 1227 GTGTTGGGAGGTTTGAGGGGTTCAGGCTCAAGTTTCGTTGGACCACTTGAAGACAGCAA 1286
 Qy 1286 GCGCGCTGCAACAGATAGGAATGATGCAAC-----CAAAATGCTTGGAGACCCCAAA 1336
 Db 1287 GAGCTCTTCAACAACCTGGGAATGATTCACATCTTCCAATAATGCTTGGAGACTCAAC 1346
 Qy 1337 GAGGTTTACCTGAAGAGCTGTCTGTCTCTTCGTTGCTGCTTTCGAGCATTTTCTTC 1396
 Db 1347 GTGGTCTCCAGAACGAGCGCTCTCAGTTCTCGGTGCTTGGCTCTTTCGAAACACTTTCTTC 1406
 Qy 1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTCTAGCAAAACGGGGCTAACAGGA 1456
 Db 1407 ATCCATACCTTAAGGATTCGGAACAGCATGCTAGCTAAGCAAAACAGGACTCACTCGTA 1466
 Qy 1457 GCCAGGTGTCTAATGCTTTCATTAATGCTCGAGTTTCGATTATGGAACCAATGGTAGAAG 1516
 Db 1467 GCCAGGTGTCTAATGCTTTCATTAATGCTCGAGTTTCGTTATGGAACCAATGGTAGG 1526
 Qy 1517 AATGTACTTGGAAAGTGAAGAAATCAAGAAACAAAACA 1555
 Db 1527 AGATGTACATGGAGAAATGAAGAGCAGGCAAGAACA 1565

RESULT 4

XX 07-OCT-2004 (first entry)
XX Cell proliferation-related nucleic acid sequence #132.
XX cell proliferation related polypeptide; cell proliferation; senescence;
XX differentiation; stress response; ds.
XX Oryza sativa.
XX W02004061122-A2.
XX 22-JUL-2004.
XX 23-DEC-2003; 2003WO-US041200.
XX 26-DEC-2002; 2002US-0436565P.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Cooper B;
XX WPI; 2004-534388/51.
XX New nucleic acid molecule encoding a cell proliferation-related
XX polypeptide, useful for modulating cell proliferation, senescence,
XX differentiation, development, and stress response in plants, and for
XX producing enhanced food crops.
XX Claim 57; SEQ ID NO 263; 408pp; English.
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a cell proliferation-related polypeptide. The nucleic acid
XX molecule and the encoded polypeptide, and methods are useful for
XX modulating cell proliferation, senescence, differentiation, development,
XX and stress response in plants, and for producing enhanced food crops. The
XX present sequence represents a cell proliferation-related nucleic acid
XX sequence. The present sequence is published separately from the main body
XX of the specification as EPO data.
XX SQ Sequence 1929 BP; 575 A; 459 C; 463 G; 432 T; 0 U; 0 Other;
Query Match 10.2%; Score 278.8; DB 12; Length 1929;
Best Local Similarity 59.0%; Pred. No. 1.2e-53;
Matches 519; Conservative 0; Mismatches 352; Indels 9; Gaps 2;
QY 765 CTCTAGTATCTGAAGCTGACAGAGCTTCTGTGATGAGTGTGTTAATATTGTTGAAA 824
DB 588 CTCAAAGTATCTGAAGCAGCACAAGAAATTGCTTGTGATGAGTGTGTTGGAAGAG 647
QY 825 AAGCATCAAAGGAGATCATCAAAGAAGGATAATTCAATCAATAAAGAATCAATGCCCTTT 884
DB 648 TATAAGCAGAGGCTCAGAAAGACCAGGCTGAAGCAGGAAATCAGATAACAAGAGC 707
QY 885 GGCTAGTGTGTCACACTAATAGTCTTGTGTTGGTGGTGAAGTAGCAGGAGGAGAAAA 944
DB 708 CGAGGGGGT-----TCGAAAGGTGAGGGGGTATCTTCCAAACCACGAGGACTACTGC 761
QY 945 TGAAGTGTGTTGAGCTTACACTGCTCAAGACAGAACTTCAATGAAGAAAAAGCAA 1004
DB 762 CAATGTGCAACCCAGAGATTTCTGCTGTGAACAACAAGAGCTCCAGAAATGAAGTGGCAA 821
QY 1005 GCTTCTTGCCATGCTTGAAGAGGTGGAGCAAGGTACAGACAGTACCATCACCAAATGCA 1064
DB 822 ACTTATGCCATGTTGATGAGGTGGACCGAAGTACAAACACTATTATCATCAATGCA 881
QY 1065 AATAATTTGTTATTCATTGAGCAAGTAGCAGGAATTTGGATCAGCCAAATCATACACTCA 1124
DB 882 AATTGTAGTCTCATCTTTTGATATGTTGCTGGGTCTGGAGCTGCCAAGCCTTATACTGC 941
QY 1125 ATTAGCTTTGATGCATTCATTCGAAGCAATTCAGATGCTTGAAGATCAATTCGTGAGCA 1184
DB 942 AGTGGCCCTTCAGACAAATCTCAAACAAATTTTCAGATGCTGAAAGATGCTATCAACGATCA 1001

QY 1185 AGTAAAGGCGACGACGAAAGAGTTTAGTGAAGAGAGGCTTGGGAGGGAATAATCGAAGG 1244
DB 1002 GATCAATGTTATCCGGAAGAACTTGGAGAGGAGGAGAGTTTCATCTGCAAGAGGGGAA 1061
QY 1245 CTCAA--GACTCAAAATTTGTGACCAATCATCTTAAGCAACAACGCGCGCTGCAACAGAT 1301
DB 1062 ATTAACGCGCGCTCCGTTATATTGACCAGCAATTAAGACAACAGCGCGCTTTCCAGCAGTA 1121
QY 1302 AGGAATGATGCAACCAAAATGCTTTGGAGACCCCAAGAGGTTTACCTCAAAAGAGCTGCTC 1361
DB 1122 TGGTTTGTACAGCAAAATGCTTTGGAGGCCACAGAGGGAGCTGCCCGAATACTCAGTTTC 1181
QY 1362 TGTCTCTTGTGCTTGGCTTTTCGAGCAATTTTCTTTCATCTTACCCTAAAGGATTTCAGACAA 1421
DB 1182 AATTCCTTGTGCTTGGCTTTTGAACACATTTCTTCCCTTCCCGGTATCCAAAGATTTCAGAAA 1241
QY 1422 AATCATGCTTGTAAAGCAACGCGGCTTAACAGGAGCCAGGTGCTTAACTGTTTCATAAA 1481
DB 1242 GCTGATGCTAGCGAGACAAATCGCTTAAACAAGTTCAGATTTCAAATTTGGTTTCATAAA 1301
QY 1482 TGCTCGAGTTTCGATTATGGAAGCCAATGTTAGAGAAATCTACTTGAAGAAGTGAAGAA 1541
DB 1302 TGCCCGTGTCCGCTGTGGAACCGATGATCGAAGACATGTATAAGAGAGATTGGGGA 1361
QY 1542 TCAAGAACAAAAACAGTACTAATACTTTCAGGAGATAACAACAAAAACAAGAGACCAATATAAG 1601
DB 1362 GCGGATCTCGACTCGAACTCTCTCCGACACGTTACCAAGGAGCAAGCAAGCAAAATAGC 1421
QY 1602 TGCTCAAAATGAAGAGAAACATCCAATTTACTTAGCAGC 1641
DB 1422 AACATCTGAGATAAGGAGATCTGAAAGCTCTATGAGC 1461
RESULT 6
AAC36460
ID AAC36460 standard; DNA; 2031 BP.
XX AC
XX AAC36460;
XX DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 13888.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 21-APR-1999; 99US-0130077P.
XX 23-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-01343768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135829P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137528P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-015263P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.

PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161933P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 9.5%; Score 261; DB 3; Length 2031;			
Best Local Similarity 66.4%; Pred. No. 1.5e-49;			
Matches 407; Conservative 0; Mismatches 200; Indels 6; Gaps 2;			
Qy	951	TGCTGTTGAGTTTCAACTCTCAAGACACAGAACTTCAATGAAAAAGCCAGCTTCT	1010
Db	1023	TACTGCTGAATCCCTCAAGCAGAGAGCGCAAGATTGCAGAGCAAGTTGTGCAAACTTTT	1082
Qy	1011	TGCCATCTTTGAAGAGGTGGGCAAGGTACAGACAGTACCATCACCAAAATGCAAAATAAT	1070
Db	1083	ATCGATATTAGCAGGTGGATAGAAATTACAAGCAGTATTACCATCAGATGCAGATAGT	1142
Qy	1071	TGTAATTATCTTTGAGCAAGTAGCAGGAATTTGGATCAGCCAAATCATACACTCAATTAGC	1130
Db	1143	TGTGTCGCTTTTCATGTAATAGCCGATGTGGAGCAGCCAAACCATACACAGCCCTTGC	1202
Qy	1131	TTTGCATGCAATTTTCAAGCAATTCAGATGCTTAAAGATGCAATTCCTGAGCAAGTAA	1190
Db	1203	GCTTCAGACCATCTCGAGGCAATTTCCGTTGCTTAAGGATGCGATATCCGACAAATCTT	1262
Qy	1191	GGCGACGACGAAGAGTTT---AGGTGAAGAGGAGGCTTGGGAGGGHAAATCGAAGGCTC	1247
Db	1263	GGTGATAAGGAAAAGTTTAGGAGGGGAACAGGATGATCAGATGGAGAGGAGTGGGGAT	1322
Qy	1248	A---AGACTCAATTTGTGACCATCATCTTAAGGCAACACGCGCGCTGCAACAGATAGG	1304
Db	1323	AAGCAGGTTAAGGAATGTTGATCAACAGTAAGGCAACAAAGAGCGTTTGCAGCGGTTAGG	1382
Qy	1305	AATGATCAACCAATCTTGGAGACCCCAAGAGGTTTACCTGAAAGAGCTGTCTCTGT	1364
Db	1383	TGTATGCAACCTCACACTTTGGCGGCTCAACGCGGTTTACCAGATTCCTCTGTTTGGT	1442
Qy	1365	CCTTCGTGCTGGCTTTTCGAGCATTTTCTTCATCTTATCCCAAGAGGATTCAGACAAAT	1424
Db	1443	TCCTCGTGTGGCTATTGAGCATTTCTCCACCTTATCCAAAGGATTCAGACAGAT	1502
Qy	1425	CATGCTTGCTAAGCAAAACGGGGCTAAACAGAGCGAGGTGTCTAACTGGTTTCATAATGC	1484
Db	1503	CATGCTCGCTAGACAAACGGGGTTGAGCCGAGGCGAGGTATCGAACTGGTTTCATAATGC	1562
Qy	1485	TCGAGTTTCGATTATGGAAGCCAAATGTTAGAGAAATCTTCTGGAAGAGTGAAGATCA	1544
Db	1563	CGGTGTCGCTCTCGGAACCGGATGGTGGAGGAGATGTACAAGGAGGAATTCAGATGC	1622
Qy	1545	AGAACAAAACAGT 1557	
Db	1623	ATTGCAGGAGAT 1635	

RESULT 7
AAC43406
ID AAC43406 standard; DNA; 1449 BP.
XX
AC AAC43406;
XX

DT	17-OCT-2000 (first entry)	
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 39140.	
XX	Hybridisation assay; genetic mapping; gene expression control;	
KW	protein identification; signal transduction pathway; metabolic pathway;	
KW	promoter; termination sequence; ss.	
XX	Arabidopsis thaliana.	
OS	EP1033405-A2.	
PN	06-SEP-2000.	
XX	25-FEB-2000; 2000EP-00301439.	
PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	28-APR-1999;	99US-0130891P.
PR	30-APR-1999;	99US-0131449P.
PR	04-MAY-1999;	99US-0132048P.
PR	05-MAY-1999;	99US-0132407P.
PR	06-MAY-1999;	99US-0132484P.
PR	07-MAY-1999;	99US-0132485P.
PR	11-MAY-1999;	99US-0132863P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	18-MAY-1999;	99US-0134370P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	25-MAY-1999;	99US-0135629P.
PR	27-MAY-1999;	99US-0136021P.
PR	28-MAY-1999;	99US-0136392P.
PR	01-JUN-1999;	99US-0136782P.
PR	03-JUN-1999;	99US-0137222P.
PR	04-JUN-1999;	99US-0137528P.
PR	07-JUN-1999;	99US-0137502P.
PR	08-JUN-1999;	99US-0137724P.
PR	10-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	14-JUN-1999;	99US-0138847P.
PR	16-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	17-JUN-1999;	99US-0139453P.
PR	18-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.

PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144325P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144331P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147302P.	Query Match 9.5%; Score 259.4; DB 3; Length 1449;		
PR	05-AUG-1999;	99US-0147192P.	Best Local Similarity 66.2%; Pred.No.3.2e-49;		
PR	06-AUG-1999;	99US-0147260P.	Matches 406; Conservative 0; Mismatches 201; Indels 6; Gaps 2;		
PR	06-AUG-1999;	99US-0147303P.	QY	951	TGCTGTTGAGCTTACAACTCTCAAGACAAAGAACTTCAAAATGAAAAAGCCCAAGCTTCT 1010
PR	06-AUG-1999;	99US-0147493P.	DB	477	TACTGCTGAAATCCCTCAAGCAGAGAGGCAAGATTGCAGAGCAAGTTGTGCGAACTTTT 536
PR	09-AUG-1999;	99US-0147935P.	QY	1011	TGCCATGCTTGAAGAGGTGGAGCAAGGTACAGACAGTACCATCACCATAATGCAAAATAT 1070
PR	10-AUG-1999;	99US-0148171P.	DB	537	ATCGATATTAGACGAGGTGGATAGAAATTACAAGCAGTAGTATTACCATCAGATGCAGATG 596
PR	11-AUG-1999;	99US-0148319P.	QY	1071	TGTTATTATCATTTGACCAAGTAGCAGGAATTGGATCAGCAAAATCATACACTCAATTAGC 1130
PR	12-AUG-1999;	99US-0148341P.	DB	597	CGTTCGCTCTTCGATGTAAATAGCCGATGTGGAGCAGCAAAACCATACACAGCCCTTGC 656
PR	13-AUG-1999;	99US-0148565P.	QY	1131	TTTGCAATGCAATTTGGAAGCAATTCAGATGCCCTAAAGAGTGCATTTGCTGAGCAAGTAA 1190
PR	13-AUG-1999;	99US-0148684P.	DB	657	GCTTCAGACCATCTCGAGGCAATTTCCGTTGCTTAAAGGATGCGATATCCGACAAATCTT 716
PR	17-AUG-1999;	99US-0149175P.	QY	1191	GGCGACGACGACAGAGTTT---AGGTGAAGAGGAGGCTTTGGGAGGGAANAATCGAAGCTC 1247
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	27-AUG-1999;	99US-0151065P.			

Db	717	GGTCGAT	AAGGNAAGTTT	TAGAGGGGAA	CAGGATGGAT	TCAGATCGG	GAGAGGAGTGGGGAT	776
Qy	1248	A---	AGACTCAAAT	TTGTGGACCA	TCACTAAGGCA	CAACACGCGCT	GCACACAGATAGG	1304
Db	777	AAGCAGGT	TAAAGAAATG	TGATCAACAGG	TAAAGGCN	CAAGACGCT	TGCAGCGTTAGG	836
Qy	1305	AATGATGA	CAACCAATG	CTTGGAGAC	CCCAAGAGG	TTTACCTGAA	AGAGCTGTCTCTGT	1364
Db	837	TGTGATGA	CAACTCAC	ACTTGGCGGC	CTCAACGCGG	TTTACCAGAT	TCTTCTGTTTGGT	896
Qy	1365	CCTTCGTG	CTTGGCTTTT	TCGAGCAAT	TTTCTTCAT	CCCTTACCC	CAANGATTTCAGACAAAT	1424
Db	897	TCTCCGTG	CTTGGCTAT	TTGAGCAT	TTCTCCAC	CCCTTATCC	AAAGATTTCAGACAAGAT	956
Qy	1425	CATGCTTGC	TAAAGCAAA	CGGGCTTAA	CAAGGAGG	CAGGTGCT	CTAACTGGTTTCATAAATGC	1484
Db	957	CATGCTCG	CTAGACAA	ACGGGGTT	GAGCGG	CGCCAGGTAT	CGNACTGGTTTCATNAATGC	1016
Qy	1485	TCGAGTT	TCGATTA	TGGAAGC	CAATGGT	TAGAAGAA	ATGTACTTGGGAAGAGTGAAGAATCA	1544
Db	1017	GCGTGTG	CTCTCTG	GAAACCG	ATGATGGT	GGAGGAGATG	TACAAGGAGGAATTCACAGATGC	1076
Qy	1545	AGAAACAAA	ACAGT	1557				
Db	1077	ATTGCAGG	AGAAAT	1089				

RESULT 8
ADG88236
ID ADG88236 standard; cDNA; 1449 BP.

RESULT 8	
ADG88236	
ID	ADG88236 standard; cDNA; 1449 BP.
XX	
XX	ADG88236;
AC	
AC	
DT	22-APR-2004 (first entry)
XX	
XX	A. thaliana RPP4-upregulated pathogen infection-related gene #578.
DE	
XX	
KW	Pathogen infection-related gene; plant; Peronospora parasitica;
KW	defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
KW	fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO200222675-A2.
PN	
XX	
PD	21-MAR-2002.
XX	
XX	14-SEP-2001; 2001WO-US028506.
XX	
XX	15-SEP-2000; 2000US-0232778P.
PR	
PR	22-JUN-2001; 2001US-0300183P.
XX	
XX	(SYGN) SYNGENTA PARTICIPATIONS AG.
PA	(UYNC-) UNIV NORTH CAROLINA.
PA	(GLAZ/) GLAZEBROOK J.
PA	(WANG/) WANG X.
PA	(DANG/) DANG J L.
PA	(EULG/) EULGEM T.
PA	(ZHUT/) ZHU T.
XX	
PI	Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
XX	
XX	WPI; 2002-292409/33.
DR	
XX	
XX	Novel isolated polynucleotide, useful for conveying pathogen resistance
PT	to plants, and for identifying plants infected with a pathogen.
PT	
XX	
XX	Claim 3; SEQ ID NO 678; 605pp; English.
XX	
CC	The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC	ADG87557)) whose expression is altered in response to pathogen infection,
CC	and to homologues of these genes from other plants or fungi, especially
CC	from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC	

RESULT 9

cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The expression of genes of the invention was upregulated or downregulated in *Arabidopsis* plants infected with the oomycete *Peronospora parasitica*, indicating that they play a role in defence mechanisms. The genes of the invention are regulated by RPP7 or RPP8 which act via unconventional signalling cascades, or by the RPP4-dependent pathway. The invention also relates to polypeptides encoded by the pathogen infection-related genes (ADG88243-ADG88327) ; expression cassettes, host cells and pathogen-resistant transgenic plants and their progeny comprising a polynucleotide of the invention; and a method of identifying a plant cell infected with a pathogen. The polynucleotide sequences and methods of the invention are useful for identifying plants infected with a pathogen, and for conferring resistance to pathogens such as oomycetes, fungi, bacteria, viruses, nematodes and insects (e.g., aphids). The present sequence represents an *Arabidopsis* thaliana gene whose expression is altered in response to *Peronospora parasitica* infection. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIFO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Query Match	9.5%	Score 259.4	DB 6	Length 1449
Best Local Similarity	66.2%	Pred. No. 3.2e-49		
Matches 406	Conservative 0	Mismatches 201	Indels 6	Gaps 2
Qy	951	TGCTGTTGAGCTTACAAC	TGCTCAAAGACAAGAACTTCAAATGAA	AAAAAGCCAAAGCTTCT 1010
Db	477	TACTGCTGAATCCCTCAAGCAGAGAGG	CNAGNAATTGCAGAGCAGAGTTGCGA	AAACTTTT 536
Qy	1011	TGCCATGCTTGAAGAGGTGGAGCAAAGGT	TACAGACAGTACCATACCAAATGCAAA	TAAAT 1070
Db	537	ATCCGATATTAGACGAGGTGGATAGAAAT	TACAAGCAGTATTACCATCAGATGCAGATAGT 596	
Qy	1071	TGTATTATCATTTTGCAAGTAGCAGGAAT	TGGATCAGCCAAATCATACACTCAAT	TAGC 1130
Db	597	CGTGTGCTCTTTTCGATGTAAATCCCGAT	TGTGGAGCAGCCAAACCATACACGCCCTTGC 656	
Qy	1131	TTTTCATGCAATTTTCAAGCAATTTCAAGT	GCCTTAAAGGATGCAATTTGCTGAGCAAGTAAA 1190	
Db	657	GCITTCAGACCATCTCGAGGCATTTCCGTT	GCTTAAGGATGCGATATCGGACAAATCTT 716	
Qy	1191	GGCGACGACGACAGATTT---AGGTCAAGAGG	AGAGCTTTGGGAGGGAATTCGAAGGCTC 1247	
Db	717	GGTGATAAGGAAAGTTTATAGAGGGGAA	CAGGATGGAATCAGATGGGAGAGGATGGGGAT 776	
Qy	1248	A---AGACTCAAAATTTGTGACCATCATCT	TAAGGCAACAAACGCGCGCTGCAACAGATAGG 1304	
Db	777	AAGCAGGTTAAGGAATGTTGATCAACAGGT	TAAGGCAACAAAGAGCGTTGCAGCGGTTAGG 836	
Qy	1305	AATGATGCACCAAAATGCTTTGGAGACCC	CAAGAGGTTTACCTGTAAGAGCTGTCTCTGT 1364	
Db	837	TGTGATGCACACCTTCACATTTGGCGGCT	CAACGCGGTTTACCAAGNTTCTTCTGTTTGGT 896	
Qy	1365	CCTTCTGCTTGCTTTTTCGAGCATTTTCTT	CACTTACCTACCCAAAGGATTCAGACAAAAT 1424	
Db	897	TCTCCGTGCTTGCTATTTGAGCATTTCTC	CAACCTTATCCAAAGGATTCAGACAAGAT 956	
Qy	1425	CATGCTTGTCTAAGCAAAACGGGGCTTAA	CAAGAGGCCAGGTGTCTAACTGGTTCTAAATGC 1484	
Db	957	CATGCTCGCTTAGACAAACGGGGTTGAG	CCGCGAGCCAGGATTCGAACTGGTTCTAATAATGC 1016	
Qy	1485	TCCAGTTTCGATTATGAAGCCCAATGTT	TAGAAGAAATGTACTTGGAAAGAGTGAAGAATCA 1544	
Db	1017	GCSTGTGCTCTCTGAAACCGATGTTG	AGGAGATGTACAGAGGAATTCACAGATGC 1076	
Qy	1545	AGAACAAAAACAGT	1557	
Db	1077	ATTGACAGGAAAT	1089	

RESULT 9

ADO62706
ID ADO62706 standard; DNA; 1971 BP.
XX
AC ADO62706;
XX
DT 15-JUL-2004 (first entry)
DE
DE Transcription factor G2550 orthologous sequence, SEQ ID 1173.
XX
XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Glycine max.
XX
XX WO2004031349-A2.
XX
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
XX
XX WPI; 2004-330163/30.
DR
XX
XX New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
XX Claim 1; SEQ ID NO 1173; 510pp; English.
XX
XX The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1971 BP; 667 A; 360 C; 451 G; 493 T; 0 U; 0 Other;
Query Match 8.9%; Score 244.4; DB 12; Length 1971;
Best Local Similarity 57.7%; Pred. No. 9.9e-46;
Matches 512; Conservative 0; Mismatches 346; Indels 30; Gaps 3;
QY 741 AAGTCTACCAACATGTTTAAAGCTCTAAGTATCTGAAGCTGCACAGAGCTTCTTGA 800
DB 201 AGCTCGAGCAACAATGTCTTGAATCTCAATACCTCAAGGAGCAGCAGAGTGTCTGA 260
QY 801 TGAAGTTGTTAATATTTGGAAAAAGCATCAAGAGATGATCAAAAAGAGATATTC 860
DB 261 TGAATAGTAAATGTCCGAAAGGCTTTGAAGCAAACTGGTTTGGAAAAAGCAAGAGTTT 320
QY 861 AATGAATA-----AAGAATCAATGCTTTGGCTAGTGTGTCAACACTAATAGTTC 911
DB 321 CGGTGACACTGGTTTAGATGGCTCCAAAGATTCTGATGGAATAATCTACAGCCAACTGT 380
QY 912 TGGTGGTGTGAAGTAGCAGCAGCAGAAAAAATGAAGTTGCTTTGAGCTTACAACTGC 971
DB 381 GCAGATGTCTTCAGGCCCAATGTTCCGCTGCTAACGCTTCTTGTGAGCTATCACCTGC 440
QY 972 TCAAGACAAAGAACTTCAATGAAAAAGCCCAAGCTTCTTGGCATGCTTGAAGAGGTGA 1031
DB 441 AGAAGCGCAGAACTTGTGTGGACAAGAAAGCAAGCTTTTGTCCATGCTGGATGAGGTGA 500
QY 1032 GCAAAGGTACAGACAGTACCATCAACCAATGCAATAAATGTTATATCATTTTGAAGCAAGT 1091
DB 501 TAAAGATACACAGCAGTACTGCCATCAGATGCAGATGTTGGTGTCTATCTTTGACATGTT 560
QY 1092 AGCAGAAATTGATCAGCAAAATCATACACTCAATTTAGCTTTGATGCAATTTGGAAGCA 1151
DB 561 TGGTGGCTGTGGAGCAGCAGAACCATATACACACTTGGCTTAAAGAAACAATTTCTCGCCA 620
QY 1152 ATTCAGATCCCTAAGAGGATGCAATTTGCTGAGCAAGTAAAGGCGCAGCAGCAAGTTTAG 1211
DB 621 CTTTGGCTGTTTGGCGTATGTCATAGTGGCCAAATTCAGGTGACCCCAAAGAGCCCTTGG 680
QY 1212 TGAAGAGGAAGGCTTGGGAGGGAATAATCGAAGGCTCAAGACTCAAAATTTGTGGACCATCA 1271
DB 681 GGAGCAAGAGG-----AATACCCCGCTCTCCGCTATGTGGATCAGCA 722
QY 1272 TCTAAGGCAACAACGCGCGCTGCAACAGATAGAAATGATGCAACCAAAATGCTTTGAGACC 1331
DB 723 ACTTAGACAACAAAAGGCCCTTCAGCAACTTGGTGTAAATGAGAC---AAGCTTTGGAGGCC 779
QY 1332 CCAAGAGGCTTACCTGAAAGAGCTGTCTCTCTTCCCTGCTTGGCTTTTCGAGCATTT 1391
DB 780 TCAGAGGGGACTTCTCTGAAACCTCTGTTTCAGTACTCCGCTGCTTGGCTCTTTGAGCATTT 839
QY 1392 TCTTCATCTCTACCCAAAGGATTTCAGACAAAATCATGCTTGTGTAAGCAAAACGGGGCTAAC 1451
DB 840 CCTTCATCTTATCTCTAAGATTTCAGAAAAATATGCTAGCAAGGCAAACTGGCTTAAC 899
QY 1452 AAGGAGCCAGGTGTCTAACTGGTTCAATAATGCTCGAGTTTCGATTTGAAAGCAATGGT 1511
DB 900 AAGAACCAAGGTGGCAAACTGGTTTCAATTAATGCAAGGGTCCGCTCTATGTAAGAGCCAAATGGT 959
QY 1512 AGAAGAAATGTTCTTGGAAAGAGTGAAGATCAAGAACAAACAGTACTAATACTTCAGG 1571
DB 960 TGAGGAAATGTTCAAAAGAAAGAAATTTGGTGAATTTCTGAGATGAGCAGCAATCTATTATCATC 1019
QY 1572 AGATAACAAAAACAAGAGACCAATATAGTGTCTCCAAATGAAGAGAA 1619
DB 1020 AGAGAAAACACTCAAGAGCTCCAAGAGATGATGTTTCAAGCTTCTGACAA 1067
RESULT 10
ABK82114

ID XX ABK82114 standard; cDNA; 706 BP.
AC AC ABK82114;
XX XX 27-AUG-2002 (first entry)
DT DT
XX XX
DE DE DNA encoding novel floral meristem identity protein LpHbb.
XX XX
XX XX Ryegrass; fescue; MADS-box; MADS; MADS-like protein; CENTRORADIALIS; CEN;
KW KW CEN-like protein; APETALA2; AP2; AP2-like protein; HB; Homeo-box protein;
KW KW HB-like protein; plant growth; plant architecture;
KW KW inflorescence development; flower development; embryo development;
KW KW seed development; flower organ identity; phase change; male sterility;
KW KW hybrid seed production; herbage quality; early maturing crop;
KW KW biomass increase; branching increase; blocking flowering;
KW KW allergenic pollen; floral meristem identity protein; gene; ss.
XX XX
XX XX Lolium perenne.
XX XX
XX XX WO200233091-A1.
XX XX
XX XX 25-APR-2002.
XX XX
XX XX 17-OCT-2001; 2001WO-AU001311.
XX XX
XX XX 19-OCT-2000; 2000AU-00000873.
XX XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
XX XX (AGRE-) AGRESEARCH LTD.
XX XX
XX XX Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;
XX XX WPI; 2002-454601/48.
XX XX P-PSDB; ABG60937.
XX XX
XX XX New substantially purified or isolated polypeptide e.g., MADS-box,
PT PT CENTRORADIALIS, APETALA2, Homeo-box proteins, isolated from ryegrass or
PT PT fescue species, useful for controlling plant life cycles and/or growth
PT PT phases.
XX XX
XX XX Claim 6; Fig 41; 290pp; English.
XX XX
CC CC The invention describes a substantially purified or isolated polypeptide
CC CC (I) from a ryegrass (*Lolium* sp.) or fescue (*Festuca* sp.) species, such as
CC CC MADS-box (MADS) and MADS-like proteins, CENTRORADIALIS (CEN) and CEN-like
CC CC proteins, APETALA2 (AP2) and AP2-like proteins, Homeo-box proteins (HB)
CC CC and HB-like proteins, or their functionally active fragments or variants.
CC CC Nucleic acid (II) encoding (I), a construct (III) comprising (II), or a
CC CC vector (IV) comprising (II), is useful for modifying plant life cycles
CC CC and/or growth phases, flowering processes, flowering and/or plant
CC CC architecture and/or flower and/or inflorescence development in a plant,
CC CC which involves introducing (II), (III) or (IV) into the plant. The
CC CC individual or simultaneous enhancement or downregulation of MADS-box gene
CC CC activities may alter flower, embryo and seed development, e.g., enhance
CC CC or inhibit embryo differentiation and growth or alter flower organ
CC CC identity through conversion of one floral organ in another. Manipulation
CC CC of CEN or HB activity in plants alters the control of phase change,
CC CC flowering time and the number of leaves made before flowering.
CC CC Manipulation of AP2 activity in plants alters flower organ identity
CC CC through conversion of one floral organ in another, leads to a change in
CC CC the number of floral organs and flowering architecture. Manipulation of
CC CC flowering plant architecture has a wide range of applications such as:
CC CC inducing male sterility for hybrid seed production; changing flower
CC CC architecture for enhancing value of ornamentals; delaying flowering in
CC CC forage grasses thus stopping the formation of less digestible stems and
CC CC increasing herbage quality; altering flowering time allowing early
CC CC maturing crops; delaying vegetative phase to increase biomass production;
CC CC increasing branching to enhanced biomass in fruit trees; altering plant
CC CC size e.g. shorter plant stature; and in blocking flowering and reducing
CC CC release for allergenic pollen. This sequence encodes a novel floral
CC CC meristem identity protein involved in floral development and a potential
CC CC target for manipulating plant life cycles
XX XX

SQ Sequence 706 BP; 210 A; 148 C; 188 G; 147 T; 0 U; 13 Other;
Query Match 8.7%; Score 237.6; DB 6; Length 706;
Best Local Similarity 66.7%; Pred. No. 2.5e-44;
Matches 355; Conservative 0; Mismatches 174; Indels 3; Gaps 1;
QY 947 AAGTTGCTGTTGAGCTTACAACCTGCTCAAGACAAGAACTTTCAAAATGAAAAAGCCCAAGC 1006
DB 173 ATGCTGTACCTGAGCTTTCACCTGCTGAGAAGCAAGAACTTCAGAACAAAGATGCCAAAC 232
QY 1007 TTCTTCCCATGCTTGAAGAGGTGGAGCAAGAGGTACAGACAGTACCACCAATGCAAA 1066
DB 233 TGATGGCAATGTTGGATGAGGTGGACCGGAAATACAAAGCACTATTACCAAAATGCAAA 292
QY 1067 TAATTGTATTATCATTTTGAACAAGTAGCAGCAATTTGATCAGGCCAAATCATACACTCAAT 1126
DB 293 ATGTGTTTCATCTTTTGTATGTTGCTGGCTGGCTGGATCTGCANAGCCGTACACTGCAG 352
QY 1127 TAGCTTTGCAATGCAATTTTGAAGCAATTCAGATGCCTTAAAGGATGCAATTTGCTGAGCAAG 1186
DB 353 TTGCTCTTCAGACAATCTCGCGCACTTCGCGTCTTGAAGGATGCCATCAATGATCAGA 412
QY 1187 TAAAGCGCAGCAGCAAGAGTTTAGT---GAAGAGGAGGCTTGGGAGGGAATCCGAG 1243
DB 413 TTAATGTTATCAGGAAGAAGCTTGGTGAGGAAGAATTCATCTGGCAAGGAAGGCAAT 472
QY 1244 GCTCAAGACTCAAAATTTGTGGACCATCATCTAAGGCAACAACGCGCTCCCAACAGATAG 1303
DB 473 TAACCCGCTCTCAGGTACATCGATCAGCAGCTGAGGCAGCAGCAGCTTCCACACAGTATG 532
QY 1304 GAATGATGCAACCAATGCTTTGAGAGACCCCAAGAGGTTTACCTGAAAAGAGCTGTCTCTG 1363
DB 533 GCATGATTCGCAAAAACGCAATGAGAGACCGCAAGAGAGACTGCTCTGAAAACCTCGGTTACGG 592
QY 1364 TCCTTCGTGCTTGGCTTTTCGAGCACTTTCTTCATCTTACCCCAAGGATTCGACACAAA 1423
DB 593 TTCTCGCTGCTTGGCTTTTCGAACACTTCTTCACCCGTACCCCAAGGATTCGAAAAGT 652
QY 1424 TCATGCTTGTAAAGCAAAACGGGCTTAACAGAGGAGCCAGGTGTCTTAAGTGGTT 1475
DB 653 TGATGCTAGGAGCAGACAGGCTTGACAAGAGCCAGATTTTCGAAATGGTT 704

RESULT 11
AAC39138
ID AAC39138 standard; DNA; 2025 BP.
XX AAC39138;
AC AAC39138;
XX XX 17-OCT-2000 (first entry)
DT DT
XX XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23512.
XX XX
XX XX Hybridisation assay; genetic mapping; gene expression control;
KW KW protein identification; signal transduction pathway; metabolic pathway;
KW KW promoter; termination sequence; ss.
XX XX
XX XX Arabidopsis thaliana.
XX XX
XX XX EP1033405-A2.
XX XX
XX XX 06-SEP-2000.
XX XX
XX XX 25-FEB-2000; 2000EP-00301439.
XX XX
XX XX 25-FEB-1999; 99US-0121825P.
XX XX 05-MAR-1999; 99US-0123180P.
XX XX 09-MAR-1999; 99US-0123548P.
XX XX 23-MAR-1999; 99US-0125788P.
XX XX 25-MAR-1999; 99US-0126264P.
XX XX 29-MAR-1999; 99US-0126785P.
XX XX 01-APR-1999; 99US-0127462P.
XX XX 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0133456P.
PR 14-MAY-1999; 99US-01334218P.
PR 14-MAY-1999; 99US-01334219P.
PR 14-MAY-1999; 99US-01334221P.
PR 14-MAY-1999; 99US-01334370P.
PR 18-MAY-1999; 99US-01334768P.
PR 19-MAY-1999; 99US-0133491P.
PR 20-MAY-1999; 99US-01335124P.
PR 21-MAY-1999; 99US-01335353P.
PR 24-MAY-1999; 99US-01335629P.
PR 25-MAY-1999; 99US-01336021P.
PR 27-MAY-1999; 99US-01336392P.
PR 28-MAY-1999; 99US-0133782P.
PR 01-JUN-1999; 99US-01337222P.
PR 03-JUN-1999; 99US-01337528P.
PR 04-JUN-1999; 99US-01337502P.
PR 07-JUN-1999; 99US-01337724P.
PR 08-JUN-1999; 99US-01338094P.
PR 10-JUN-1999; 99US-01338540P.
PR 10-JUN-1999; 99US-01338847P.
PR 14-JUN-1999; 99US-01339119P.
PR 16-JUN-1999; 99US-01339452P.
PR 16-JUN-1999; 99US-01339453P.
PR 17-JUN-1999; 99US-01339492P.
PR 18-JUN-1999; 99US-01339454P.
PR 18-JUN-1999; 99US-01339455P.
PR 18-JUN-1999; 99US-01339456P.
PR 18-JUN-1999; 99US-01339457P.
PR 18-JUN-1999; 99US-01339458P.
PR 18-JUN-1999; 99US-01339459P.
PR 18-JUN-1999; 99US-01339460P.
PR 18-JUN-1999; 99US-01339461P.
PR 18-JUN-1999; 99US-01339462P.
PR 18-JUN-1999; 99US-01339463P.
PR 18-JUN-1999; 99US-01339750P.
PR 21-JUN-1999; 99US-01339817P.
PR 22-JUN-1999; 99US-01339899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 13-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.

```
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159323P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 8.6%; Score 234.6; DB 3; Length 2025;
Best Local Similarity 58.6%; Pred. No. 1.8e-43;
Matches 480; Conservative 0; Mismatches 309; Indels 30; Gaps 3;

QY 739 GGAAGTTCTAGCAACATGGTTTATAGGCTCTAAGTATCTGAAGTCTGACAGAGCTTCTT 798
DB 669 GGAGGGTTTATGAGCAGCGTTCTCGGTCTCTCGGTATCTTAACTCTGCTCAGATTGCTT 728
QY 799 GATGAAGTTGTTAATATTGTTGGAAAGCATCAAGGAGATGATCAAAAGAGATTAAT 858
DB 729 GATGAAGTGTGTAGTGTCAAGAAAGAACTAAACCAATGATGAAGAAAGATGAA--- 784
QY 859 TCAATGAATAAGAAATCAATGCTTTGGCTAGTGTCTCAACTAATAGTCTTGGTGGT 918
DB 785 -----AGTTAATGACTTTTAACTATGTTCTTAGGATAGAAGGAGGAGGCA 831
QY 919 GGTGAAGTAGTACGACGAGCAGAGAAATGAAGTTGCTGTGAGCTTACAACTGCTCAAGA 978
DB 832 GTGCGAGCTATCGAATG-ATTCGAATGGGAATCGATTGAGTTATCTACTATTGAACGT 890
QY 979 CAAGAACTTCAATGAATAAAGCAACGCTTCTGCCATGTTGAGAGGTGGACAAAG 1038
DB 891 GATGAGCTTCAATAAAGAAAGAACAGCTTTTAACTAATGGTTGATGAGGTAGATAAAGA 950
QY 1039 TACAGACAGTACCATCAACCAATGCAAAATGAATGTTATTCATTGAGCAAGTAGCAGGA 1098
DB 951 TATAACCAATATTACCAATCAATGAAGAGCATGACTTCATCAATTTGAGATAGTAGCAGGA 1010
QY 1099 ATTGGATCAGCAATCATACACTCAATTAGCTTTGCAATTTGCAAGCAATTCAGA 1158
DB 1011 CTTGGATCAGTAAAGCTTACACATCAGTTGCTCTCAACAGAACTCTCGCCATTTTCGT 1070
QY 1159 TGCTTAAAGATGCAATTTGCTGACCAAGTAAAGCGCAGCAGCAAGATTTAGTGAAGAG 1218
DB 1071 GCTCTTCGCGACGCAATAAAGGACAAATTCAGATTGTTAGAGAGAACTTGGGGAGAAA 1130
QY 1219 GAAGGCTTGGAGGGAATTCGAGGC-----TCAAGACTCAAAATTTGTGGAC 1266
DB 1131 GGAGGAGATGCTTGGATGAGCAACAAGGAGAGGAGATCAAGACTCAGGTATTTAGAT 1190
QY 1267 CATCATCTAAGGCAACACGCGCTGCAACAGATAGGAATGATGCAACCAAAATGCTTGG 1326
DB 1191 CAACGGTTGAGACAGCAAGAGCTTTGATCAACAGCTTGAATGGTTGACCGCTTGG 1250
QY 1327 AGACCCCAAGAGGTTTACCTGAAAGAGCTGTCTCTGCTCTCTGCTGCTGCTTTCGAG 1386
DB 1327 AGACCCCAAGAGGTTTACCTGAAAGAGCTGTCTCTGCTCTCTGCTGCTGCTTTCGAG 1386
```

```
DB 1251 AGACCTCAAAAGAGTCTTCTGTAATACTCTGTTTCTGTTCTCGTGTCTGGCTTTCGAA 1310
QY 1387 CATTTTCTTCATCCTTACCCAAAGGATTTCAGACAAAATCATGTTGCTTAAGCAAAACGGG 1446
DB 1311 CATTTCTTCATCATATCGAAGAAATCTGAGAAATCATGTTGCGAAACACACAGGA 1370
QY 1447 CTAAAGAGCCAGGTGTTCTAATGTTTCAATGTTTCAATGTTGCTGAGTTTCGATTATGGAAGCCA 1506
DB 1371 CTGTCGAAGAACCCAGGTGCTAATTTGTTTCTAATAACGCGAGAGTTCGTTATGGAACCG 1430
QY 1507 ATGCTAGAGAAATGTTACTTGGGAAGTGAAGTGAAGTCAAA 1545
DB 1431 ATGATCGAGAGATGTTAATAAGAGAGTTTGGTGATGAA 1469

RESULT 12
AAD06493
ID AAD06493 standard; cDNA; 1575 BP.
XX
AC AAD06493;
XX
DT 10-AUG-2001 (first entry)
XX
DE Arabidopsis thaliana G418 transcription factor homologue, G2550 cDNA.
XX
KW Transcription factor; pesticidal; antimicrobial; gene therapy;
KW pathogen tolerance; trichome structure; callose induction;
KW phytoalexin induction; plant structure; plant development; ss.
XX
OS Arabidopsis thaliana.
XX
Key Location/Qualifiers
CDS 1..1575
FT /tag= a
FT /product= "A. thaliana transcription factor homologue"
XX
PN WO200135726-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-US031418.
XX
PR 17-NOV-1999; 99US-0166228P.
PR 17-APR-2000; 2000US-0197899P.
PR 22-AUG-2000; 2000US-0227439P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (HEAR/) HEARD J.
PA (RATC/) RATCLIFFE O.
PA (CREE/) CREELMAN R.
PA (JIAN/) JIANG C.
PA (PINE/) PINEDA O.
PA (REUB/) REUBER L.
PA (ADAM/) ADAM L.
XX
PI Heard J, Ratcliffe O, Creelman R, Jiang C, Pineda O, Reuber L;
PI Adam L;
XX
PI WPI; 2001-335978/35.
XX P-PSDB; AAE02524.
XX
PT Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the pathogen resistance characteristics of plants, e.g.
PT corn, potato and cotton plants.
XX
XX Claim 4; Page 110-112; 134pp; English.
XX
XX The present sequence is a cDNA encoding Arabidopsis thaliana
XX transcription factor homologue. The transcription factors are used to
XX alter the structure and developmental characteristics of plants such as
XX soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,
XX alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry,
XX raspberry, carrot, cantaloupe, cauliflower, cucumber, coffee, eggplant,
```

CC grapes, mango, lettuce, honeydew, melon, onion, papaya, peppers,
CC pineapple, spinach, squash, sweet corn, tobacco, tomato, peas,
CC watermelon, rosaceous fruits and vegetable brassicas. The transcription
CC factors are specifically useful for modifying traits associated with
CC plant's pathogen tolerance such as alterations in cell wall composition,
CC trichome number or structure, callose induction, phytoalexin induction,
CC and alterations in the cell death response. Transgenic plants expressing
CC these transcription factors are more tolerant to biotrophic or
CC necrotrophic pathogens such as fungi, bacteria, molluscs, viruses,
CC nematodes and parasitic higher plants. The transcription factors are also
CC used in gene therapy
XX
SQ Sequence 1575 BP; 512 A; 323 C; 323 G; 417 T; 0 U; 0 Other;

Matches	392	Conservative	0	Mismatches	241	Indels	3	Gaps	1
Qy	958	GAGCTTAA	CAATCTGCTCA	AAAGACA	AGAACTCTTCA	AAATGAAAAAGCA	AGCTTCTTTGGCAATG	1017	
Db	670	GAGTTGTCTCTCT	TTTCAGAACGCTCA	GAGAGCTAC	GAGCAAGCAAGCA	AGCTTTTAAACAATG	729		
Qy	1018	CTTCG	AGAGCTGGAGCA	AAAGGTAC	AGACAGTACC	ATCACC	AAATGCAAAATTTGTTATTA	1077	
Db	730	GTGGAT	GAGGTAGATAA	AAAGGTATA	CCCAATAC	CATCATCAAT	TGGAAGCTTTAGCATCG	789	
Qy	1078	TCAATTT	GAGCAAGTAG	CAGGAAATGG	GATCAG	CCAAATCATAC	ACTCAATTAGCTTTGTCAT	1137	
Db	790	TCATTT	GAGATGGTAA	CAGGTCTTGG	AGCAGCTTA	AGCCTTTAC	ACATCCGTAGCTCTGAAT	849	
Qy	1138	GCAATTT	CGAAGCAATTC	AGATGCCTTA	AGAGTGC	CAATTCCT	GAGCAGNAGTAAAGGCGAGC	1197	
Db	850	AGAACTCTCT	CGCCATTTTC	CGCTGTTTAC	CGCAGC	CGGATAA	AAAGAACAGATTCAGGTGATC	909	
Qy	1198	AGCAAGAGTTT	TAGGTGA	AGAGGAGGCTTC	GGGAGG	GGAAAAATCGA	AGG- --CTCAAGACTC	1254	
Db	910	AGAGGGA	AGCTTGGG	AGAGAGAGACTTCT	GTATG	ATGAC	AGGAGAGAGATACCGCGTCTT	969	
Qy	1255	AAATTTGTG	GACCAATCATCTTA	AGGCACAA	ACGGCGCTG	CAACAGATAG	AGAAATGATGCAAA	1314	
Db	970	AGGTACTT	TAGATCAAC	CGGTTTG	AGACAA	CAGAGAGCTTTG	GCATCAACAACCTTGGAAATGGTT	1929	
Qy	1315	CCAAA	TGCTTTGG	AGACCCCA	AGAGGTTTAC	CTTGAA	AGAGCTGTCTCTGTCTTCTCGTGCT	1374	
Db	1030	AGACCAGCTT	GGAGACCA	AGAGGCTTAC	CTG	AAAACTCTGTCTCTA	TACTTCTGAGCT	1089	
Qy	1375	TGGCTTTT	TCGAGCATTTTCT	TCATCCTTACC	CAAAAGGATTC	CAGACAAAATCAT	CGCTTGCT	1434	
Db	1090	TGGCTCTTT	GAGCATTTTCTT	TCATCC	ATATCTTAA	AGAAATCAG	AGAAAAATCATGCTTTCA	1149	
Qy	1435	AAGCAAC	CGGGCTTA	CAGAGGCC	AGGTGTCTTAA	CTGGTTTCAT	TAATGTCTCAGATTCGA	1494	
Db	1150	AAGCAGAC	AGGACTAT	CGAAAAAC	CGGTTG	CAAAATTTGGTTTATTA	ACGCGAGAGTTTCGA	1209	
Qy	1495	TTATGGA	AGCCAAATGGT	PAGAAGAAATGTA	CTTGG	AGAAGTGA	AGAAATCAAGAACCAAAAC	1554	
Db	1210	CTATG	GAACCAATGATG	GAAGATGTATA	AGAAGAGGTTT	GGAGNAATCAG	CAGAGTTA	1269	
Qy	1555	AGTACTA	ATATCTT	CAGGAGATA	CAAAAA	CAAAAGAG	1590		
Db	1270	CTCTCTA	ACTCTAAT	CAAGAC	CAAAAAA	AAATG	CGAG1305		

	CC	content, altered seed protein content, altered zeaxanthin content,
	CC	altered leaf prenyl lipid content, increased anthocyanin levels, and
	CC	decreased anthocyanin levels. Note: The sequence data for this patent
	CC	not form part of the printed specification, but was obtained in
	CC	electronic format directly from WIPO at
	CC	ftp.wipo.int/pub/published_pct_sequences.
	XX	
	SQ	Sequence 1575 BP; 512 A; 323 C; 323 G; 417 T; 0 U; 0 Other;
RESULT 13		
ADO61968		
ID	ADO61968 standard; DNA; 1575 BP.	
XX		
AC	ADO61968;	
XX		
DT	15-JUL-2004 (first entry)	
XX		
DE	Transcription factor G2550 coding sequence, SEQ ID 435.	
XX		
XX		
	Query Match	8.6%; Score 234.4; DB 12; Length 1575;
	Best Local Similarity	61.6%; Pred. No. 1.8e+43;

KW	Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW	osmotic stress tolerance; cold tolerance; heat tolerance;
KW	low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW	glyphosate resistance; flowering; fertility; seed development; ds.
XX	
OS	<i>Arabidopsis thaliana</i> .
XX	
XX	WO2004031349-A2.
XX	
PD	15-APR-2004.
XX	
XX	18-SEP-2003; 2003WO-US030292.
XX	
XX	18-SEP-2002; 2002US-0411837P.
PR	17-DEC-2002; 2002US-0434166P.
PR	24-APR-2003; 2003US-0465809P.
XX	
XX	(MEND-) MENDEL BIOTECHNOLOGY INC.
PA	
XX	Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI	Riechmann JB, Haake V, Dubell AN, Keddle JS, Sherman BK;
PI	
XX	WPI; 2004-330163/30.
DR	P-PSDB; ADO61969.
DR	
XX	
PT	New recombinant polynucleotide encoding transcription factor
PT	polypeptides, useful for producing transgenic plants with advantageous
PT	properties compared to a reference plant.
XX	
PS	Claim 1; SEQ ID NO 435; 510bp; English.

Query Match 8.6%; Score 234.4; DB 12; Length 1575;
Best Local Similarity 61.6%; Pred. No. 1.8e-43;

Matches 392; Conservative 0; Mismatches 241; Indels 3; Gaps 1;		
Qy	958 GAGCTTCAACTGTCTCAAGACAAGAACTTCAAAATGAAAAGCAAGCTTCTTGCCATG 1017	PR 25-MAR-1999;
Db		PR 29-MAR-1999;
Qy	670 GAGTTGTCTCTCAGAACGTCAGAGCTACAGAGCAAGAGCTTTTAACAAATG 729	PR 01-APR-1999;
Db		PR 06-APR-1999;
Qy	1018 CTTGAAGAGGTGGAGCAAAAGGTA CAGACAGTACCATCACCAAAATGCAAAATAATTTGTAATTA 1077	PR 08-APR-1999;
Db		PR 16-APR-1999;
Qy	730 GTGGATGAGGTAGATAAAAGGTATAACCAATACCATTCAATGGAAGCTTTAGCATCG 789	PR 19-APR-1999;
Db		PR 21-APR-1999;
Qy	1078 TCATTTGAGCAAGTAGCAGGAATTTGGATCGCCCAATATACACTCAATTAGCTTTGCAT 1137	PR 23-APR-1999;
Db		PR 28-APR-1999;
Qy	790 TCITTCGATGGTAAACAGGCTTTGGAGCAGCTTAAGCCTTACACATCCGTAGCTCTGAAT 849	PR 30-APR-1999;
Db		PR 04-MAY-1999;
Qy	1138 GCAATTTGGAAGCAATTCAGATGCTTAAAGGATGCAATTTGCTGAGCAAGTAAAGCGCAGC 1197	PR 05-MAY-1999;
Db		PR 06-MAY-1999;
Qy	850 AGAATCTCTGCCATTTCCGCTGTTTACGCGACGGATAAAGAACAGATTCAGTGTATC 909	PR 07-MAY-1999;
Db		PR 11-MAY-1999;
Qy	1198 AGCAAGAGTTTAGTGTAAGAGGAAGGCTTTGGGAGGGGAAAATCGAAGG---CTCAAGACTC 1254	PR 14-MAY-1999;
Db		PR 14-MAY-1999;
Qy	910 AGAGGGAAGCTTTGGGGAGAGAGAGACTTCTGATGAACAAGGAGAGAGATACCGGTCTT 969	PR 14-MAY-1999;
Db		PR 14-MAY-1999;
Qy	1255 AAATTTGTGACCATCATCTAAAGCAACAACGCGCTGCAACAGATAGGAAATGATGCAA 1314	PR 14-MAY-1999;
Db		PR 18-MAY-1999;
Qy	970 AGGTACTTAGATCAACGGTTGAGACACAGAGAGCTTTGCAATCAACAACCTTGGAAATGGTT 1029	PR 19-MAY-1999;
Db		PR 20-MAY-1999;
Qy	1315 CCAATGCTTTGGAGACCCCAAGAGGTTTACTTGAAGAGCTGTCTCTGTCCTTTCGTGCT 1374	PR 21-MAY-1999;
Db		PR 24-MAY-1999;
Qy	1030 AGACCAGCTTTGGAGACCAACAAGAGGCTTACCTGAAAACCTGTCTCTATACTTCGAGCT 1089	PR 25-MAY-1999;
Db		PR 27-MAY-1999;
Qy	1375 TGGCTTTTCGAGCAATTTCTTCATCTTACCCAAAGGATTCAGACAAAATCATGCTTTGCT 1434	PR 28-MAY-1999;
Db		PR 01-JUN-1999;
Qy	1090 TGGCTCTTTGAGCAATTTCTTTCATCTCATATCTTAAAGAAATCAGAGAAAATCATGCTTTCA 1149	PR 03-JUN-1999;
Db		PR 04-JUN-1999;
Qy	1435 AAGCAACGGGGCTAAACAAGGAGCCAGGTGTCTAACTGGTTTATAATGCTCGAGTTGCA 1494	PR 07-JUN-1999;
Db		PR 08-JUN-1999;
Qy	1150 AAGCAGACAGGACTATCGAAAACACAGGTTCGCAAAATTTGGTTTATTAACGGCAGAGTTGCA 1209	PR 10-JUN-1999;
Db		PR 10-JUN-1999;
Qy	1495 TTATGGAAGCCAAATGGTAGAAGAAATGTACTTGAAGAAGTGAAGAATCAAGAACAAAAC 1554	PR 14-JUN-1999;
Db		PR 16-JUN-1999;
Qy	1210 CTATGGAACCAATGATTGAAGATGTATAAAGAAAGATTGGAGAAATCAGCAGATTGA 1269	PR 16-JUN-1999;
Db		PR 17-JUN-1999;
Qy	1555 AGTACTAATCTTCAGGAGATAACAAAACAAAAGAG 1590	PR 18-JUN-1999;
Db		PR 18-JUN-1999;
Qy	1270 CTCTCTAACTCTAATCAAGACACCAACCAAAAAAATGCAG 1305	PR 18-JUN-1999;
Db		PR 18-JUN-1999;
RESULT 14		
AAC36745		
ID	AAC36745 standard; DNA; 1824 BP.	
AC	AAC36745;	
XX		
XX		
DT	17-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 14930.	
XX		
XX		
KW	Hybridisation assay; genetic mapping; gene expression control;	
KW	protein identification; signal transduction pathway; metabolic pathway;	
KW	promoter; termination sequence; ss.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
XX	25-FEB-2000; 2000EP-00301439.	
PR		
PR	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.	
PR	09-MAR-1999; 99US-0123548P.	
PR	23-MAR-1999; 99US-0125788P.	

PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	28-APR-1999;	99US-0130891P.
PR	30-APR-1999;	99US-0131449P.
PR	04-MAY-1999;	99US-0132048P.
PR	05-MAY-1999;	99US-0132407P.
PR	06-MAY-1999;	99US-0132485P.
PR	07-MAY-1999;	99US-0132486P.
PR	11-MAY-1999;	99US-0132487P.
PR	14-MAY-1999;	99US-0132863P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139452P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	24-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-01444085P.

```
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 05-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 28-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 01-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 03-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.

PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 8.6%; Score 234.4; DB 3; Length 1824;
Best Local Similarity 61.6%; Pred. No. 1.9e-43;
Matches 392; Conservative 0; Mismatches 241; Indels 3; Gaps 1;

QY 958 GAGCTTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAGCCAAAGCTTCTTGGCAATG 1017
Db 697 GAGTTGTCTCTTCAGAACGTCAGGAGCTACAGACGAAGAAGCAAGCTTTTAAACAATG 756
QY 1018 CTTGAAGAGGTGGAGCAAAAGGTACACAGACAGTACCATCACCAAATGCAAAATGTTATTA 1077
Db 757 GTGGATGAGGTAGATAAAAGGTATAACCAATACCAATCAATCAAAATGGAAGCTTTAGCATCG 816
QY 1078 TCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGTCAT 1137
Db 817 TCTTCGAGATGTTAAACAGGCTTTGGAGCAGCTTAAGCCTTACACATCCGTAAGTCTGAAT 876
QY 1138 GCAATTTGGAAGCAATTCAGATGCGCTTAAAGGATGCAATTCAGCAAGTAAAGGGGACG 1197
Db 877 AGAATCTCTCGCCATTTCCGCTGTTTACGCGACGCGATATAAAGAACACAGATTTCAGGTGATC 936
QY 1198 AGCAAGAGTTTAGTGAAGAGGAGGCTTGGGAGGGAATTCGAAG--CTCAAGACTC 1254
Db 937 AGAGGGAAGCTTGGGGAGAGAGAGACTTCTGATGAACAAGAGGAGGATACCGCGTCTT 996
QY 1255 AATTTTGTGGACCATCATCTAAGGCAACAACGCGCTGCAACAGATAGGAATGATGCAA 1314
Db 997 AGGTACTTAGATCAACGGTTGAGACAACAGAGAGCTTTGCATCAACAACCTTGGATGGTT 1056
QY 1315 CCAATGCTTTGGAGAGCCCAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCTGTCGT 1374
Db 1057 AGACCAGCTTGGAGACCACAAGAGGCTTACCTGAAAACTCTGCTCTATCTACTTCCAGCT 1116
QY 1375 TGGCTTTTCGAGCATTTTCTTCATCTTACCCAAAGGATTCACAGAAAATCATGCTTGCT 1434
Db 1117 TGGCTCTTTGAGCATTTCTTTTCATCATATTCCTTAAAGAAATCAGAGAAAATCATGCTTCA 1176
QY 1435 AAGCAAAACGGGCTAAACAAGAGCAGGCTGCTTAACTGGTTCTATAAATGCTCGAGTTTCA 1494
Db 1177 AAGCAGACAGGACTATCGAAAAACAGGTTGCAAAATGGTTTATTAACGCGAGAGTTTCA 1236
QY 1495 TTATGGAAGCCAAATGTTAGAGAAATGTTACTTGGAGAGTGAAGTAATCAAGAACAAAC 1554
Db 1177 AAGCAGACAGGACTATCGAAAAACAGGTTGCAAAATGGTTTATTAACGCGAGAGTTTCA 1236
QY 1495 TTATGGAAGCCAAATGTTAGAGAAATGTTACTTGGAGAGTGAAGTAATCAAGAACAAAC 1554
Db 1177 AAGCAGACAGGACTATCGAAAAACAGGTTGCAAAATGGTTTATTAACGCGAGAGTTTCA 1236
```

Db 1237 CTATGGAACCAATGATTGAAGAGATGTATTAAGAGAGTTTGGAGAAATCAGCAGAGTTA 1296
Qy 1555 AGTACTAATCTTCAGGAGATAACAAAACAAAGAG 1590
Db 1297 CTCTTAATCTAATCAAGACACCAAAAATGATGAG 1332

RESULT 15
AD062707 standard; DNA; 2458 BP.
XX AC AD062707;
XX 15-JUL-2004 (first entry)
XX Transcription factor G2550 orthologous sequence, SEQ ID 1174.
XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW phosphate resistance; flowering; fertility; seed development; ds.
XX Glycine max.
XX WO2004031349-A2.
XX 15-APR-2004.
XX 18-SEP-2003; 2003WO-US030292.
XX 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddie JS, Sherman BK;
XX WPI; 2004-330163/30.
XX New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX Claim 1; SEQ ID NO 1174; 510pp; English.
XX The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (AD061534-AD063778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased

CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenly lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 2458 BP; 751 A; 488 C; 545 G; 674 T; 0 U; 0 Other;
Query Match 8.5%; Score 231.8; DB 12; Length 2458;
Best Local Similarity 56.8%; Pred. No. 8.4e-43;
Matches 481; Conservative 0; Mismatches 342; Indels 24; Gaps 2;
Qy 741 AAGTTCTAGCAACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGA 800
Db 779 ACGTTACTCAACAGATATTTTGAACCTCTCAATACCTGAAGCGCGCACAGATTTCGTTGA 838
Qy 801 TGAAGTTGTTAATATTGTTGAAAAAGCATCAAGGAGATGATCAAAAGAGGTAATTC 860
Db 839 TGAATAGGTAGTGCGCGAAAGGCTTTGAAGCAATCTGGAATGGAAAAACAAGAGAACAC 898
Qy 861 AATGAATAAAGAAATCAATGCTTTTGGCTAGTGTGTCAACACATAATAGTTCTGGTGGTGG 920
Db 899 TGGATTAGATGGCTCTAAGATTTCTGATGGAATAATCTACAGTCATCTATGCAAAATGTC 958
Qy 921 TGAAGTAGCAGCAGCAGCAAAAAATGAAGTTGCTGTTTGAAGCTTTACAACTGCTCAAGACA 980
Db 959 TTCAGGCCCTTAATGGTTCCTCACTGCTAATGCTTCTAGTGAACATATCATCCGACAGCGCA 1018
Qy 981 AGAATCTCAATGAAAAAGCCCAAGCTTCTTGGCATGCTTGAAGAGCTGGAGCAAAAGTA 1040
Db 1019 GAATCTGTTGGACAAGAAAAACAAGCTTTTGTCCATGTTGGATGAGGTAGATAAAAGATA 1078
Qy 1041 CAGACAGTACCATCACCMAATGCAAAATAATTTGTTATTATATTTTGGAGCAAGTAGCAGGAAT 1100
Db 1079 CAGACAGTACTGCCATCAGATGAGATGTTGGTGTCTATCTTTTGACATGGTTGCTGCTG 1138
Qy 1101 TGGATCAGCCAAATCATACACTCAATTTAGCTTTTGCATGCAATTTTGAAGCAATTTAGATG 1160
Db 1139 TGGAGCAGCAGAACCATATACACACTTGCCTTTAAGAAACAATTTCTGCCACTTTCCGTG 1198
Qy 1161 CCTAAGGATGCAATTTGCTGAGCAAGTAAAGCGCAGCAGCAAGATTTAGGTGAAGAGGA 1220
Db 1199 TTTGCGTGTAGCCATCAGTGGCCAAATTCAGGTGAGCCCAAGAGCCCTTGGGGAGCA --- 1255
Qy 1221 AGGCTTGGGAGGGAAAAATCGAAGGCTCAAGACTCAAAATTTGTGGACCATCATCTTAAGGCA 1280
Db 1256 -----AGGAATACCCCGTCTCCGCTATGTGGATCAGCAACTTAGACA 1297
Qy 1281 ACAACGGCGCTGCAACAGATAGGAATGTATGCAACCAAAATGCTTTGGAGAGACCCCAAGAGG 1340
Db 1298 ACAAAAGGCACCTTCAGCAACTTGGTGTATAGAGAC---AAGCTTGGAGGCTCTCAGAGGGG 1354
Qy 1341 TTTACCTGAAGAGCTGTC 1400
Db 1355 TCTTCTGAAAGCTCTGTTTCAATATCTCCGTGCTGGCTCTTTTGAACATTTTCTTCATCC 1414
Qy 1401 TTACCCAAAGGATTCAGACAAAAATCATGCTTGTCTAAGCAAAACGGGGCTAAACAAGAGCCA 1460
Db 1415 TTACCCCAAGGATTCGAGAAAAATTTATGCTGGCTAGGCAAAACAGGCTTTGACCAGAACCA 1474
Qy 1461 GGTGTTCTAATCTGTTTCAATAATGCTCGAGTTTCGATTTATGAAAGCAATTTAGTAGAGAAAT 1520
Db 1475 GGTGGCCAAATTTGTTTCATCAATGCAAGGGTGCCTCTCTGGAAGCCAAATTTGTTAGGAAAT 1534
Qy 1521 GTACTTGGAGAGAGTGAAGATCAAGAACAAAAACAGTACTTAATCTTCAGAGATACAA 1580
Db 1535 GTACAAAGAGGAGTTTGGAGACTCCGAGATGAAACTCCAACTTTCATCTCTGAGAACAAAC 1594

QY 1581 AAACAAA 1587
|||
Db 1595 GGTCAAA 1601

Search completed: July 29, 2005, 08:33:46
Job time : 1419 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 07:15:42 ; Search time 445 Seconds
(without alignments)
10056.667 Million cell updates/sec

Title: US-10-624-201A-1

Perfect score: 2735
Sequence: 1 catcgagagataaaatata.....gcaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	183.8	6.7	529	4	US-09-640-211A-77
2	184.6	6.0	343	4	US-09-640-211A-1578
3	159.8	5.8	323	4	US-09-640-211A-418
4	155.8	5.7	316	4	US-09-640-211A-1854
5	155.8	5.7	462	4	US-09-640-211A-359
6	136.2	5.0	1239	4	US-09-640-211A-65
7	134.2	4.9	381	4	US-09-640-211A-102
8	134.2	4.9	381	4	US-09-640-211A-1472
9	121	4.4	380	4	US-09-640-211A-111
10	121	4.4	380	4	US-09-640-211A-1246
11	120	4.4	690	4	US-09-640-211A-400
12	111.4	4.1	698	4	US-09-640-211A-37
13	107.4	3.9	366	4	US-09-640-211A-1241
14	105.4	3.9	474	4	US-09-640-211A-57
15	105.4	3.9	474	4	US-09-640-211A-1399
16	96	3.5	171	4	US-09-640-211A-1845
17	89.4	3.3	158	4	US-09-640-211A-370
18	84.6	3.1	260	4	US-09-640-211A-351
19	84.6	3.1	260	4	US-09-640-211A-383
20	84.2	3.1	157	4	US-09-640-211A-394
21	81.6	3.0	486	4	US-09-640-211A-50
22	60.6	2.2	533	4	US-09-640-211A-185
23	59.8	2.2	30820	4	US-09-949-016-17145
24	59.4	2.2	407	4	US-09-640-211A-383
25	59.2	2.2	7218	1	US-08-232-463-14
26	58.2	2.1	18773	4	US-09-949-016-14164
27	56.4	2.1	19124	2	US-08-487-826B-13
Sequence 13, Appl					
Sequence 14, Appl					
Sequence 15, Appl					
Sequence 16, Appl					
Sequence 17, Appl					
Sequence 18, Appl					
Sequence 19, Appl					
Sequence 20, Appl					
Sequence 21, Appl					
Sequence 22, Appl					
Sequence 23, Appl					
Sequence 24, Appl					
Sequence 25, Appl					
Sequence 26, Appl					
Sequence 27, Appl					
Sequence 28, Appl					
Sequence 29, Appl					
Sequence 30, Appl					
Sequence 31, Appl					
Sequence 32, Appl					
Sequence 33, Appl					
Sequence 34, Appl					
Sequence 35, Appl					
Sequence 36, Appl					
Sequence 37, Appl					
Sequence 38, Appl					
Sequence 39, Appl					
Sequence 40, Appl					
Sequence 41, Appl					
Sequence 42, Appl					
Sequence 43, Appl					
Sequence 44, Appl					
Sequence 45, Appl					

Sequence 1357, Ap
Sequence 5219, Ap
Sequence 648, App
Sequence 12776, A
Sequence 15940, A
Sequence 1031, Ap
Sequence 381, App
Sequence 17145, A
Sequence 13703, A
Sequence 2156, Ap
Sequence 12378, A
Sequence 22, Appl
Sequence 195, App
Sequence 12699, A
Sequence 168053,
Sequence 168054,
Sequence 168055,

ALIGNMENTS

RESULT 1

US-09-640-211A-77
; Sequence 77, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; US-09-640-211A-77

Query Match 6.7%; Score 183.8; DB 4; Length 529;
Best Local Similarity 77.0%; Pred. No. 2.6e-37;
Matches 224; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 1258 TTTGTGGACCATCATCTTAAGGCAACACGGCGCTGCAACAGATAGGAATGATCAACA 1317
Db 38 TTTTGGACCAAGAGCTCAGGCAACAGAGCCCTACAGCAGCTTGGGATGATCAACAG 97
Qy 1318 AATGCTGGAGACCCCAAGAGGTTTACCTGAAGAGCTGTCTCTGCTCTGCTGG 1377
Db 98 CATGATGGAGGCGCGCAAGAGGACTTCTTGAGAGTTCTGTCTATTTCTTGGGCGCTGG 157
Qy 1378 CTTTTCGAGCATTTTCTTCTATCCTTACCAAGAGATTGAGCAAAATCATCTGCTTAAG 1437
Db 158 CTATTTGAGCATTTTCTTCTATCCTTACCAAGAGATTGAGCAAAATCATCTGCTTAAG 217
Qy 1438 CAAACGGGGCTTAACAGAGGCCAGGTGTCTAACTGGTTCTAAATGCTCGAGTTCGATTA 1497
Db 218 CAGACAGGCTTGAAGAAGTCTCAGGTCTCGAATTTGGTTTCATCAATGCAAGAGTGGCTCTC 277
Qy 1498 TGGAGCCATGGTAGAAGAAATGACTTGGAGAGAGTGAAGATCAAGAA 1548
Db 278 TGGAAACCTATGTCGAGAAATGTACAAAGAGAGATTTGGGGATGCGGAA 328

RESULT 2

US-09-640-211A-1578
; Sequence 1578, Application US/09640211A

```

; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Sherk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
;      Transcription of Gene
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1578
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Pinus radiata
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(343)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-211A-1578

```

```

RESULT 3
US-09-640-211A-418
; Sequence 418, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shengk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 418
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-418

```

Qy	1284	ACGGCGCTCGAACAGATAGGAATGATGCAACCAATATGCTTGGAGAGACCCCAAGAGAGTTT	1343
Db	1	ACGGGCTCTCCAAACAATTAGGCATGATTCAGCAGCATGCTTGGAGGCCACAGAGAGGACT	60
Qy	1344	ACCTGAAAGAGCTGCTCTCTGCTCTTGGCTTTCGGCTTTTCGAGCATTTTCTTCATCCTTA	1403
Db	61	TCCGAGCGATCTGTTTCTCTCTTACCGGCTTGGCTATTTGAACATTTTCTTCATCCGTA	120
Qy	1404	CCCAAGGATTACAGACAAAATCATGCTTGTGTAAAGCAAAACGGGGCTAACAGGAGCCAGGT	1463
Db	121	TCCAAAAGATGCGACAGCAAAACATATGCTCGCGACAGACACTGGGCTTACCAGAAATCAGGT	180
Qy	1464	GTCTACTGGTTCATTAATGCTCGAGTTCGATTATGGAAGCCCAATGTTAGAGAAATGTA	1523
Db	181	CTCAATTGGTTTATAATTCAGGTGTACGCTCTTGGAGCCCTATGGTGGAGAGATGTA	240
Qy	1524	CTTGGAGAAAGTGAAGAATCAAGAACAAAAC	1554
Db	241	TGTGGAGGAAACAAAGGAGCGCAGAGTAGAC	271

[illegible]

```
Db 252 GCTGATAAGCATCTGTGGCTGCACAGACTGGTCTCTCCAGAAACCAGGTCTCGAATTGG 311
Qy 1474 TTCTATAATGCTCGAGTTCGATTATGGAAGCCCAATGGTAGAAGAAATGTACTTTGGAGA 1532
Db 312 TTCTATAATGCCAGGTTCGGTTGTGGAAACCCATGGTGAGGAGATGTACCAGCAAGA 370

RESULT 8
US-09-640-211A-1472
; Sequence 1472, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1472
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1472

Query Match 4.9%; Score 134.2; DB 4; Length 381;
Best Local Similarity 65.6%; Pred. No. 1.4e-24;
Matches 196; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 1234 AAAATGGAAGGCTCAAGACTCAAAATTTGTGGACCATCATTAAGGCAACACCGCGCTG 1293
Db 72 AAGGGCGAGACACACCGGCTCAAGTTGCTGCAGCAGAGCTGTAGGCGAGCAGAGGGCTTTC 131
Qy 1294 CAACAGATAGGAATGATGCAACCAAAATGTTGGAGACCCCAAGAGGTTTACCTGAAAGA 1353
Db 132 CACAGATGGGCATGATGAGCAAGAGCCTGGAGCGCGAGCGGCTGCCGAGCGG 191
Qy 1354 GCTGTCTCTGCTTCGTGCTTGGCTTTTTCGAGCATTTTCTTCATCTTACCCAAAGGAT 1413
Db 192 TCGGTCAACATACCTGCGTCATGGCTCTTCGAGCATTTCTTCGATCCGTATCCAGTGAC 251
Qy 1414 TCAGACAAATCATGCTTGTGTAAGCAACGGGCTCAACAGGAGCCAGGTGTCTACTGG 1473
Db 252 GCTGATAAGCATCTGTGGCTGCACAGACTGGTCTCTCCAGAAACCAGGTCTCGAATTGG 311
Qy 1474 TTCTATAATGCTCGAGTTCGATTATGGAAGCCCAATGGTAGAAGAAATGTACTTTGGAGA 1532
Db 312 TTCTATAATGCCAGGTTCGGTTGTGGAAACCCATGGTGAGGAGATGTACCAGCAAGA 370

RESULT 9
US-09-640-211A-111
; Sequence 111, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-111

Query Match 4.4%; Score 121; DB 4; Length 380;
Best Local Similarity 80.2%; Pred. No. 3.4e-21;
Matches 142; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1377 GCTTTTCGAGCATTTTCTTCATCTCTTACCCAAAGGATTCAGACAAATCATGCTTGGCTAA 1436
Db 1 GCTCTTCGAACACTTTTCTCCACCCCTACCCGAGGATTCGACAAAGTTCATGCTGGCCAA 60
Qy 1437 GCAAAACGGGGCTAAACAAGGAGCCAGGTGTCTAACTGGTTTCATAAAATGCTCGAGTTCGATT 1496
Db 61 ACAGACAGGGCTCACTAGAAAGCCAGGTGTCTGAATTTGTTTATAAAATGCTCGAGTTCGGCT 120
Qy 1497 ATGGAAGCCCAATGGTAGAAGAAATGTACTTTGGAAGAGTGAAGAATCAAGAACAATA 1553
Db 121 TTGGAAGCCGATGGTGGAGGAGATGTACACGGAGGAAATCAAGGACCAAGAACAGAA 177

RESULT 10
US-09-640-211A-1246
; Sequence 1246, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1246
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1246

Query Match 4.4%; Score 121; DB 4; Length 380;
Best Local Similarity 80.2%; Pred. No. 3.4e-21;
Matches 142; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1377 GCTTTTCGAGCATTTTCTTCATCTCTTACCCAAAGGATTCAGACAAATCATGCTTGGCTAA 1436
Db 1 GCTCTTCGAACACTTTTCTCCACCCCTACCCGAGGATTCGACAAAGTTCATGCTGGCCAA 60
Qy 1437 GCAAAACGGGGCTAAACAAGGAGCCAGGTGTCTAACTGGTTTCATAAAATGCTCGAGTTCGATT 1496
Db 61 ACAGACAGGGCTCACTAGAAAGCCAGGTGTCTGAATTTGTTTATAAAATGCTCGAGTTCGGCT 120
Qy 1497 ATGGAAGCCCAATGGTAGAAGAAATGTACTTTGGAAGAGTGAAGAATCAAGAACAATA 1553
Db 121 TTGGAAGCCGATGGTGGAGGAGATGTACACGGAGGAAATCAAGGACCAAGAACAGAA 177

RESULT 11
US-09-640-211A-400
; Sequence 400, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
```

```
; ORGANISM: Eucalyptus grandis
US-09-640-211A-111

Query Match 4.4%; Score 121; DB 4; Length 380;
Best Local Similarity 80.2%; Pred. No. 3.4e-21;
Matches 142; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1377 GCTTTTCGAGCATTTTCTTCATCTCTTACCCAAAGGATTCAGACAAATCATGCTTGGCTAA 1436
Db 1 GCTCTTCGAACACTTTTCTCCACCCCTACCCGAGGATTCGACAAAGTTCATGCTGGCCAA 60
Qy 1437 GCAAAACGGGGCTAAACAAGGAGCCAGGTGTCTAACTGGTTTCATAAAATGCTCGAGTTCGATT 1496
Db 61 ACAGACAGGGCTCACTAGAAAGCCAGGTGTCTGAATTTGTTTATAAAATGCTCGAGTTCGGCT 120
Qy 1497 ATGGAAGCCCAATGGTAGAAGAAATGTACTTTGGAAGAGTGAAGAATCAAGAACAATA 1553
Db 121 TTGGAAGCCGATGGTGGAGGAGATGTACACGGAGGAAATCAAGGACCAAGAACAGAA 177

RESULT 10
US-09-640-211A-1246
; Sequence 1246, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1246
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1246

Query Match 4.4%; Score 121; DB 4; Length 380;
Best Local Similarity 80.2%; Pred. No. 3.4e-21;
Matches 142; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1377 GCTTTTCGAGCATTTTCTTCATCTCTTACCCAAAGGATTCAGACAAATCATGCTTGGCTAA 1436
Db 1 GCTCTTCGAACACTTTTCTCCACCCCTACCCGAGGATTCGACAAAGTTCATGCTGGCCAA 60
Qy 1437 GCAAAACGGGGCTAAACAAGGAGCCAGGTGTCTAACTGGTTTCATAAAATGCTCGAGTTCGATT 1496
Db 61 ACAGACAGGGCTCACTAGAAAGCCAGGTGTCTGAATTTGTTTATAAAATGCTCGAGTTCGGCT 120
Qy 1497 ATGGAAGCCCAATGGTAGAAGAAATGTACTTTGGAAGAGTGAAGAATCAAGAACAATA 1553
Db 121 TTGGAAGCCGATGGTGGAGGAGATGTACACGGAGGAAATCAAGGACCAAGAACAGAA 177

RESULT 11
US-09-640-211A-400
; Sequence 400, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
```



```

; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 400
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-400

Query Match      4.4%; Score 120; DB 4; Length 690;
Best Local Similarity 78.3%; Pred. No. 8.4e-21;
Matches 144; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 939 GAAAAATGAAGTCTCTGCTGAGCTTACAACTGCTCAAGACAGCAAGCACTTCAATGAAAAA 998
Db 504 GAAAAAGCAGCATGCTTTTGCACTCACACCAGCTGATAGACAGCAAGCACTTCAATGAAAAA 563
Qy 999 AGCCAAGCTTCTGCCATGCTTGAAGAGGTGGAGCAAGGTACAGACAGTACCATCACCA 1058
Db 564 GGCAGAGCTTGTGCCATGTTGGATGAGGTGGATCGAGGTACAGACAGTACTATCATCA 623
Qy 1059 AATGCAAAATATGTTATATCATTTGAGCAAGTATGAGGAAATGGATCAGCCAAATCATA 1118
Db 624 GATGCAAAATCTGTTTTCATGCTTTGAGACCGCAGCTGGATTTGGGGCTGCCAAGACATA 683
Qy 1119 CACT 1122
Db 684 CACT 687

RESULT 12
US-09-640-211A-37
; Sequence 37, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 698
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-37

Query Match      4.1%; Score 111.4; DB 4; Length 698;
Best Local Similarity 63.0%; Pred. No. 1.4e-18;
Matches 172; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 953 CTGTTGAGCTTACAACTGCTCAAGACAGCAAGCACTTCAATGAAAAAGCCAAAGCTTTG 1012
Db 415 CTAGCGAGCTTCTGCTGCTGAAAAACAAGATTTCAGACAAAGCTCACGAAGCTCTTGT 474
Qy 1013 CCATGCTTGAAGAGTGGAGCAAGGTACAGACAGTACCATCCCAATGCAATTAATTTG 1072
Db 475 CCATGTTGGACGAGGTTGATAAAGGTACAGCAGTACTATCACCAATGCAGATCGTGG 534
Qy 1073 TATTATCATTTGACAGTACGAGGAATTTGGATCAGCCAAATCATACACTCAATTAGCTT 1132
Db 535 TACAGTCTTTTGATACATAGCAGGAGCGGTGAGCCAGCCCTACACGCGCTTGGC 594
Qy 1133 TGCATGCAATTTGAAAGCAATTCAGATGCTTAAAGGATGCAATTTGTCGACAGCAAGTAAAG 1192
Db 595 TCCAGAGGATATCCCGCACTTCCGGTCCCTGTCATGACGGGATCACGGGTCAAATTCAG 654
Qy 1193 CGACGACAGAGTTTATAGGTGAAGGCAAGGCT 1225

```

```

Db 655 CAACCCGTAAAAAGTCTCGGAGAGCAAGACACCT 687

RESULT 13
US-09-640-211A-1241
; Sequence 1241, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1241
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1241

Query Match      3.9%; Score 107.4; DB 4; Length 366;
Best Local Similarity 78.2%; Pred. No. 1.1e-17;
Matches 129; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 1384 GAGCATTTTCTTCATCTTACCCAAAGGATTCAGACAAATCATGCTTGAAGCAACG 1443
Db 1 GAGCATTTTCTTCATCTTACCCAAAGGATTCAGACAAATCATGCTTGAAGCAACG 60
Qy 1444 GGGCTAACAGAGGAGCAGGTGCTAACTGGTTCTATAAATGCTCGAGTTCCGATTATGAAG 1503
Db 61 GGCTTCACAGAGTCAAGTCTCGAATTTGTTTCATCATGCAAGAGTGGCTCTCTGAAA 120
Qy 1504 CCAATGGTAGAAGAAATGTACTTTGGAAGAGTGAAGAAATCAAGAA 1548
Db 121 CCTATGTCGAAGAAATGTACAAAGAGAGATTGGGATCGCGAA 165

RESULT 14
US-09-640-211A-57
; Sequence 57, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-57

Query Match      3.9%; Score 105.4; DB 4; Length 474;
Best Local Similarity 77.9%; Pred. No. 4e-17;
Matches 127; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 1402 TACCCAAAGGATTCAGACAAATCATGCTTGTAAAGCAACCGGGCTTAAACAGGAGCCAG 1461
Db 126 TATCCCAAGATTTCGACAAACACATGCTCGCAAAACAGCGGAGTAAACAGGAGCCAG 185
Qy 1462 GTGTCTAACTGGTTTCATAAATGCTCGAGTTTCGATTATGGAAGCCCAATGGTAGAAGAAATG 1521

```

```
Db      186 GTCTCTAACTGGTTTCATCAACGCTCGGGTTCGCTCTGGAAGCCCATGGTCTGAAGAAATG 245
Qy      1522 TACTTGGAGAAGTGAAGATCAAGAACAAAACAGTACTAATA 1564
Db      246 TACTTGAAGAGACCAAGAGCCGAGAGCAAGCTGGGTCTGAGA 288
```

```
RESULT 15
US-09-640-211A-1399/c
; Sequence 1399, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1399
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1399
```

```
Query Match      3.9%; Score 105.4; DB 4; Length 474;
Best Local Similarity 77.9%; Pred. No. 4e-17;
Matches 127; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy      1402 TACCCAAAGGATTGAGACAAAATCATGCTTGCTTAAGCAACGGGCTTAACAGGCCAG 1461
Db      349 TATCCCAAAGATTTCGGACAAACACATGCTCGCAAAACAGCGGGACTAACCCAGGAGCCAG 290

Qy      1462 GTGCTTAACCTGGTTTCATTAATGCTCGAGTTCGATTATGGAGCCAAATGGTAGAAGAAATG 1521
Db      289 GTGCTTAACCTGGTTTCATCAACGCTCGGGTTCGCTCTGGAAGCCCATGGTCTGAAGAAATG 230

Qy      1522 TACTTGAAGAAGTGAAGATCAAGAACAAAACAGTACTAATA 1564
Db      229 TACTTGAAGAGACCAAGAGCCGAGAGCAAGCTGGGTCTGAGA 187
```

Search completed: July 29, 2005, 14:20:30
Job time : 449 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 08:10:18 ; Search time 1673 Seconds
(without alignments)
10580.930 Million cell updates/sec

Title: US-10-624-201A-1
Perfect score: 2735
Sequence: 1 catcgagagataaaatata.....gcaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	432.4	15.8	3033	18	US-10-424-599-130628, Sequence 130628,
2	392	14.3	1555	18	US-10-425-114-13353, Sequence 13353, A
3	378.2	13.8	1606	18	US-10-425-114-7804, Sequence 7804, Ap
4	353	12.9	1822	18	US-10-424-599-133558, Sequence 133558,
5	352.6	12.9	1892	18	US-10-424-599-63874, Sequence 63874, A
6	346.8	12.7	3311	18	US-10-424-599-141176, Sequence 141176,
7	324.6	11.9	2043	9	US-09-938-842A-1337, Ap

8	324.6	11.9	2043	11	US-09-938-842A-1337	Sequence 1337, Ap
9	324.6	11.9	2043	21	US-10-495-918-153	Sequence 153, App
10	324.6	11.9	2385	17	US-10-225-066A-335	Sequence 335, App
11	324.6	11.9	2385	17	US-10-374-780A-2671	Sequence 2671, Ap
12	324.6	11.9	2385	22	US-10-225-066A-335	Sequence 335, App
13	314	11.5	1716	18	US-10-425-114-9882	Sequence 9882, Ap
14	304	11.1	582	18	US-10-424-599-7028	Sequence 7028, Ap
15	295.6	10.8	3028	19	US-10-437-963-41007	Sequence 41007, A
16	278.8	10.2	2799	19	US-10-437-963-31585	Sequence 31585, A
17	268.8	9.8	1649	18	US-10-425-114-20970	Sequence 20970, A
18	266	9.7	677	19	US-10-767-701-3524	Sequence 3524, Ap
19	265.6	9.7	2943	20	US-10-425-115-118586	Sequence 118586,
20	264	9.7	831	19	US-10-767-701-10301	Sequence 10301, A
21	257.4	9.4	3029	19	US-10-437-963-60154	Sequence 60154, A
22	245.6	9.0	1645	18	US-10-424-599-65908	Sequence 65908, A
23	244.2	8.9	1637	18	US-10-425-114-13359	Sequence 13359, A
24	243.2	8.9	1546	18	US-10-425-114-29785	Sequence 29785, A
25	243.2	8.9	2736	18	US-10-424-599-37574	Sequence 37574, A
26	240	8.8	1498	18	US-10-425-114-33974	Sequence 33974, A
27	240	8.8	2313	18	US-10-425-114-15195	Sequence 15195, A
28	240	8.8	2970	20	US-10-425-115-97705	Sequence 97705, A
29	230.2	8.4	1452	18	US-10-425-114-5047	Sequence 5047, Ap
30	227	8.3	1452	20	US-10-425-115-118580	Sequence 118580,
31	226.2	8.3	1280	18	US-10-425-114-13349	Sequence 13349, A
32	225.4	8.2	1599	9	US-09-938-842A-2154	Sequence 2154, Ap
33	225.4	8.2	2486	18	US-10-424-599-55280	Sequence 55280, A
34	225.4	8.2	1884	9	US-09-938-842A-905	Sequence 905, App
35	224	8.2	1884	11	US-09-938-842A-905	Sequence 905, App
36	224	8.2	1983	14	US-10-286-264-17	Sequence 17, Appl
37	224	8.2	1983	17	US-10-225-066A-819	Sequence 819, App
38	224	8.2	1983	17	US-10-374-780A-2481	Sequence 2481, Ap
39	224	8.2	1983	17	US-10-412-699B-637	Sequence 637, App
40	224	8.2	1983	22	US-10-225-066A-819	Sequence 819, App
41	224	8.1	2545	17	US-10-225-068-235	Sequence 235, App
42	222.6	8.1	2545	17	US-10-374-780A-2329	Sequence 2329, Ap
43	222.6	8.1	2545	18	US-10-412-699B-267	Sequence 267, App
44	222.6	8.1	2545	21	US-10-225-068-235	Sequence 235, App

ALIGNMENTS

RESULT 1

US-10-424-599-130628
; Sequence 130628, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 130628
; LENGTH: 3033
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_88966C.1
US-10-424-599-130628

Query Match 15.8%; Score 432.4; DB 18; Length 3033;
Best Local Similarity 72.4%; Pred. No. 3.8e-90;
Matches 575; Conservative 0; Mismatches 216; Indels 3; Gaps 1;
Qy 758 TTTTAGCTCTAGTATCTCAAGCTGCACAGAGCTTCTTGATGAAGTTGTTAATTG 817
Db 861 TTATGGGTCAAGTACTTGAAGCTGCACAGAGCTTCTTGATGAAGTTGTTAATTG 920

```
QY 818 TTGMAAAGCATCAAGAGAGATGATCAAAAGAGGATATTCATGAATAAAGAAATCAA 877
Db 921 GAAAGGAAATCTCAAAAGAGAGAAATCTCAGAGAGCGTGAAGCGAATAGGGAATCAA 980
QY 878 TGCCCTTTGGCTAGTGTGTCAACACATAATATGTTCTGGTGGTGTGAAAGTAGCAGCAGGC 937
Db 981 CTAATTCGTGCTGCTGGTGATGCTGGAGATGGTTCAAGTGGTGGAGGAGAAACAGTG 1040
QY 938 AGAAATGAAGTTGCTGTGTGAGCTTCAACTGCTCAAGACAAAGAACTTCAAAATGAAA 997
Db 1041 CAGGGAACAAGT---GGTTGAATCAGCACCGCTCAGAGGCAAGAGCTTCAGATGAAGA 1097
QY 998 AAGCAGCTTCTTGCCACTGCTTGAAGAGGTGGAGCAAGGTTACAGACGTACCATCAC 1057
Db 1098 AGTCCAGCTTGTGACCATGCTCGATGAGGTAGAACAAAGGTACAGCAATATCACCA 1157
QY 1058 AAATGCAAAATAATTTGTTATATCAATTTGAGCAAGTAGCAGGAAATTTGATCAGCCAAATCAT 1117
Db 1158 AAATGCAAAATTTGGTATCGTCAATTTGAGCAAGCAGCAGGTATTTGGGCGGCAAAATCTT 1217
QY 1118 AACTCAATTAGCTTTGCATGCAATTTGGAAGCAATTCAGATGCTTAAAGGATGCAATG 1177
Db 1218 ACATGCCCCCTGGCCCTAAAAAATATCTCAAAGCAATTCAGGTGCTGAAAGATGCAATCT 1277
QY 1178 CTGAGCAAGTAAAGGGCAGCAGCAAGAGTTTGTGTAAGAGGAGGCTTTGGGAGGGA 1237
Db 1278 CTGCACAAATCAAGCCACAGCAAGACATTTGGGTGAAGATGATGCTTAGGAGTTAAAG 1337
QY 1238 TCGAAGGCTCAAGACTCAAAATTTGTGACCATCATCTTAAGGCAACACGCGCTCGAAC 1297
Db 1338 TAGAAGGCTCAAGGCTTAGGTATGTTGACCAACCATCTCAGGCAACAAAGGGCACTTCAGC 1397
QY 1298 AGATGAATATGATCAACCAATGCTTGGAGACCCCAAGAGGTTTACCTGAAAGAGCTG 1357
Db 1398 AGCTAGGAATGATTCACCAATGATGAGGAGGCCCCAAAGAGGCTTGCCTGAAAGAGCTG 1457
QY 1358 TCTCTGCTTCTGCTGCTGCTTTGAGGATTTTCTGAGCATTTTCTTATCCCAAGAGTTTCA 1417
Db 1458 TTTCCATCTTCGGGCTTGGCTTTTGGAGCATTTTCTTACCTTATCCCAAGAGCTTCG 1517
QY 1418 ACAAAATCATGCTTGTGTAAGCAACCGGGCTTAACAGGAGCCAGGTGTCTTAACCTGTTCA 1477
Db 1518 ATAAAGTTATGCTTGTGTAACAAACTGGAATGCTAGGAGCCAGGTGTCAAACTGGTTTA 1577
QY 1478 TAAATGCTCAGTTCGATTATGGAAGCAATGTTAGAGAAATGTTACTTGGAGAGAGTGA 1537
Db 1578 TCAATGCCCGAGTTCGGCTTTTGGAGCCAAATGTTGAAAGAAATGTACTTGGAGAAATCA 1637
QY 1538 AGAATCAAGAACAA 1551
Db 1638 AGGAGCATGAACAA 1651
```

```
RESULT 2
US-10-425-114-13353
; Sequence 13353, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; LENGTH: 1555
```

```
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMST02400042B10_FLI
US-10-425-114-13353
Query Match 14.3%; Score 392; DB 18; Length 1555;
Best Local Similarity 78.4%; Pred. No. 7.4e-81;
Matches 488; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 955 GTTGAGCTTACAACTGCTCAAGACAAAGAACTTCAAAATGAAAAGAACCCAAAGCTTCTTGCC 1014
Db 92 GTTGAATCTCAGCACCGCTCAGAGGCAAGAGCTTCAGATGAAGAACTTGTGACC 151
QY 1015 ATGCTTTGAGAGGTGAGCAAGGTACAGACAGTACCATCACCAGTACCAATGCAAAATGTA 1074
Db 152 ATGCTCGATGAGGTAGAACAAAGGTACAGCAATATCACCCCAAAATGCAAAATTTGGTA 211
QY 1075 TTATCAATTTGAGCAAGTAGCAGGAAATTTGATCAGCCAAATCATACACTCAATTTAGCTTTG 1134
Db 212 TCGTCAATTTGAGCAGCAGCAGGTTATGGGCGGCAAAATCTTACACTGCCCTTGCCTTA 271
QY 1135 CATGCAATTTGGAAGCAATTCAGATGCTTAAAGGATGCAATTTGCTGAGCAAGTAAAGCG 1194
Db 272 AAGACTATCTCAAAGCAATTTGAGGTGCTTAAAGATGCAATCTCTGCACAAAATCAAAGCC 331
QY 1195 ACGAGCAAGAGTTTAGTGAAGAGGAGGCTTTGGAGGGGAAATCGAAGGCTCAAGACTC 1254
Db 332 ACGAGCAAGACATTTGGGTGAAGATGATTTTAGGAGTTAAAGTAGAAGGCTCGAGCTT 391
QY 1255 AAATTTGTGGACCATCATCTTAAGGCAACAAACGCGCTCGCAACAGATAGGAATGATGCAA 1314
Db 392 AGTGTGTTGACCAACCATCTCAGGCAACAAAGGGCACTTCAGCAGCTAGGAATGATCAA 451
QY 1315 CCAATGCTTGGAGACCCCAAGAGGTTTACCTGAAAGAGCTGTCTCTGCTTCTGCTGCT 1374
Db 452 CCNATGCTTGGAGGCCCCAAAGAGGCTTGCCTTGAACGAGCTGTTTCCATTTCTTCGGCT 511
QY 1375 TGCTTTTTCGAGCATTTTCTCATCTTACCCCAAGGATTCAGACAAATCATGCTTGTCT 1434
Db 512 TGCTTTTTCGAGCATTTTCTTACCTTATCCCTTATCCAAAGAGCTCCGATAAAGTTATGCTTGT 571
QY 1435 AAGCAACCGGGCTTAACAAAGGAGCCAGGTGTCTTAACCTGTTTCAATAATGCTCGAGTTGCA 1494
Db 572 AAACAACTGGACTTCTAGGAGCCAGGTGTCAAACTGGTTTATCAATGCCCGAGTTGCG 631
QY 1495 TTATGAGACCAATGTTAGAGAAATGTTACTTGGAGAGAGTGAAGATCAAGAACAA 1551
Db 632 CTTTGAAGCCATGTTTGAAGAAATGTACTTGGAGAAATCAAGGAGCATGAACAA 688
```

```
RESULT 3
US-10-425-114-7804
; Sequence 7804, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7804
; LENGTH: 1606
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
```

OTHER INFORMATION: Clone ID: 700684286_FLI
US-10-425-114-7804

Query Match 13.8%; Score 378.2; DB 18; Length 1606;
Best Local Similarity 73.6%; Pred. No. 1.3e-77;
Matches 482; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

Qy 908 GTTCTGGTGGTGAAGTAGCAGCAGGCGAGAGAAATGAAGTTGCTGTTGAGCTTACAA 967
Db 10 GTGAGGAGGAGGAGAAATAATAATATGATGAGGCAACAAGGAGTTGAACTCAGCA 69

Qy 968 CTGCTCAAAGACAGAACTTCAATGAAAAGCCAGCTTCTTGCCATGCTTGAAGAGG 1027
Db 70 CAGCACAGACAGAGAGCTTCAGATGAAGAAGTCAAAACTTTGTACCATGCTAGATGAGG 129

Qy 1028 TGGAGCAAGGTACAGACAGTACCATCACAAATGCAAAATTAATGTTATTCATTTCGAGC 1087
Db 130 TGGAGCAAGGTACCGACAGTATCACCAAAATGCAAGTTGATGAATCAATTCATTTCGAGC 189

Qy 1088 AAGTAGCAGGAATTTGGATAGCCAAATCATACACTCAATTAAGTTTGCATGCAATTTGGA 1147
Db 190 AAGCAGCGGGTGTGGAGCTGCAAAAGTCTTACACAGCCCTTGCCTTAAAGACAATCTCAA 249

Qy 1148 AGCAATTCAGATGCTTAAAGATGCAATTTGCTGAGCAAGTAAAGCGCAGGCAAGAGTT 1207
Db 250 AGCAATTCGGTGTCTCAAGATGCAATCTCTTCAAAATTAAGACCAAGCAGGCAAAACCT 309

Qy 1208 TAGGTGAAGGAGGAGGCTTGGAGGGGAAATCGAAGGCTCAAGACTCAAAATTTCTGGACC 1267
Db 310 TAGGTGAAGATGATGTTCTTGGAGTTAAGGTAGAGGTTTCGAGGCTTAGGTATGTTGATC 369

Qy 1268 ATCATCTAAGGCAACAAGCGCGCTGCAACAGATAGGAATGATGCAACCAAAATGCTTGA 1327
Db 370 ATCAACTGAGACAACAAGTGCATTAACAACAGCTTGGAAATGATCAACAACAAATGCTTGA 429

Qy 1328 GACCCCAAGAGGTTTACCTGAAAGAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1387
Db 430 GGGCCCAAGAGGCTTGCCTGAAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 489

Qy 1388 ATTTCTTCATCTCTTACCAGGATTCAGACAAATCATCTGCTTGTCTGCTTGTGCTGCTGCTG 1447
Db 490 ATTTCTTGCACTCTTATCTTAAAGACTCCGATAAGTTATGCTTGCAAAACAACTGAC 549

Qy 1448 TAACAAGGAGCAGGTTGCTTAACCTGTTCAATAATGCTCGAGTTGATTTGGAAGCCAA 1507
Db 550 TTACTCGAGCAGGTTGCTTAACCTGTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609

Qy 1508 TGGTAGAAGAAATGCTACTTGAAGAGTGAAGATCAAGACAAACAGTACTAA 1562
Db 610 TGGTAGAAGAAATGCTACTTGAAGAGGTAAGAAACAAAGAACCAATAGCTCTCA 664

RESULT 4
US-10-424-599-133558
; Sequence 133558, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 133558
; LENGTH: 1822
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91608C.1

US-10-424-599-133558

Query Match 12.9%; Score 353; DB 18; Length 1822;
Best Local Similarity 72.6%; Pred. No. 1.1e-71;
Matches 471; Conservative 0; Mismatches 175; Indels 3; Gaps 1;

Qy 908 GTTCTGGTGGTGAAGTAGCAGCAGGCGAGAGAAATGAAGTTGCTGTTGAGCTTACAA 967
Db 19 GTGAGGAGGAGGAGAAATAATAATATGATGAGGCAACAAGGAGTTGAACTCAGCA 78

Qy 968 CTGCTCAAAGACAGAACTTCAATGAAAAGCCAGCTTCTTGCCATGCTTGAAGAGG 1027
Db 79 CAGCACAGACAGAGAGCTTCAGATGAAGAAGTCAAAACTTTGTACCATGCTAGATGAGG 138

Qy 1028 TGGAGCAAGGTACAGACAGTACCATCACAAATGCAAAATTAATGTTATTCATTTCGAGC 1087
Db 139 TGGAGCAAGGAGGAGCAGTATCACCAAAATGCAAGTTGATGAATCAATTCATTTCGAGC 198

Qy 1088 AAGTAGCAGGAATTTGGATAGCCAAATCATACACTCAATTAAGTTTGCATGCAATTTGGA 1147
Db 199 AAGCAGCGGGTGTGGAGCTGCAAAAGTCTTACACAGCCCTTGCCTTAAAGACAATCTCAA 258

Qy 1148 AGCAATTCAGATGCTTAAAGATGCAATTTGCTGAGCAAGTAAAGCGCAGGCAAGAGTT 1207
Db 259 AGCAATTCGGTGTCTCAAGATGCAATCTCTTCAAAATCAAGACAACTAGCAAGACCT 318

Qy 1208 TAGGTGAAGGAGGAGGCTTGGAGGGGAAATCGAAGGCTCAAGACTCAAAATTTCTGGACC 1267
Db 319 TAGGTGAAGATGATGTTCTTGGAGTTAAGGTAGAGGTTTCGAGGCTTAGGTATGTTGATC 378

Qy 1268 ATCATCTAAGGCAACAAGCGCGCTGCAACAGATAGGAATGATGCAACCAAAATGCTTGA 1327
Db 379 ATCAACAGACAGACAACAAGTGCATTAACAACAGCTTGGAAATGATCAACAACAACTTGA 435

Qy 1328 GACCCCAAGAGGTTTACCTGAAAGAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1387
Db 436 GGGCCCAAGAGGCTTGCCTGAAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495

Qy 1388 ATTTCTTCATCTCTTACCAGGATTCAGACAAATCATCTGCTTGTCTGCTTGTGCTGCTGCTG 1447
Db 496 ATTTCTTGCACTCTTATCTTAAAGACTCCGATAAGTTATGCTTGCAAAACAACTGGC 555

Qy 1448 TAACAAGGAGCAGGTTGCTTAACTGTTCAATAATGCTCGAGTTGATTTGGAAGCCAA 1507
Db 556 TTACTCGAGCAGGTTGCTTAACTGTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615

Qy 1508 TGGTAGAAGAAATGCTACTTGAAGAGTGAAGATCAAGACAAACAG 1556
Db 616 TGGTAGAAGAAATGCTACTTGAAGAGGTAAGAAACAAAGAGCCAAACAAATAG 664

RESULT 5
US-10-424-599-63874
; Sequence 63874, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 63874
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28692C.1
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28692C.1
US-10-424-599-63874

; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1337
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1337

Query Match 11.9%; Score 324.6; DB 9; Length 2043;
Best Local Similarity 65.2%; Pred. No. 5.1e-65;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

```
Qy 758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAGAGCTTCTTGATGAAGTTGTTAATATTG 817
Db 569 TTGTAGCTCCAACTACTTGAAGCGCACACAGAGCTTCTTGACGAAGTAGTCAACGCTG 628
Qy 818 TTGGAAGAGCTCAAGAGGAGATGATCAAAAGAGGATTAATCAATGAATAAAGAAATCAA 877
Db 629 ATTCCGATGACATGACGCTAATCCCACTATTCTCATCGAAAAGGGTAGTTGCGGAA 688
Qy 878 TGCCTTTGGCTAGTGTCAACACTAATATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 937
Db 689 ATGATAAAACCTGTCGAGAAATCATCGCGCGCTGCGAGGAGAGGTTCCGGTGGCGAG 748
Qy 938 AGAAAAATG-----AGTTGCTGTTGAGCTTACACTGCTCAAGAGCAAGAACTTCAAA 991
Db 749 CAGAAGCAGCGCGGAAACGTCGCTGAGCTTAGGCACGCGCAGAGAGACAAGAAATACAGA 808
Qy 992 TGAAGGAGGCAAGCTTCTTCCCATGCTTGAAGAGGTGCGAGCAAGAGGTACAGACAGTACC 1051
Db 809 TGAAGGAGGCAAACTTAGTAACTATGCTTCAAGAGGTGAGAGAGATATAGACAGTACC 868
Qy 1052 ATCAACAAATGCAAAATATGTTATATCATTTGAGCAAGTAGCAGGAATGGATCAGCCA 1111
Db 869 ACCAGCAGATGCAGATGCTGATCTCTTCGTCGAGCAAGCGCAGGAGTAGGATCAGCGA 928
Qy 1112 AATCATACACTAATTTGCTATGCTCAATTTTGAAGCAATTCAGATGCTTAAGAGATG 1171
Db 929 AGTCATACACGTCGCTAGCATTTGAATCCATATCAAGACAGTTCCTGCTTGAAGAGG 988
Qy 1172 CAATTGCTGAGCAAGTAAAGCGCAGCAGAGAGTCTTAGGTGAAGAGGAGGCTTG--- 1227
Db 989 CGATCGCTGCTAGATTAAGCGGCAACAGAGTCTTGGGAGGAGATTCAGTGTCTG 1048
Qy 1228 --GGAGGGAATAACGAGGCTCAAGCTCAAAATTTTGGACCATCATCTAAGGCAACAA 1285
Db 1049 GTGTTGGAGGTTTGAAGGCTCAGGCTCAAGTTCGTTGGACCACTTGGAGAGAGG 1108
Qy 1286 GCGGCTGCAACAGATTCAGACAAATCATGCTTGAAGCAAAACGCGGCTAACAGGA 1336
Db 1109 GAGCTCTTCAACAACTCGGAATGATTCACATCCTTCCAAATATGCTTGGAGACCTCAAC 1168
Qy 1337 GAGGTTTACCTGAAAGAGTGTCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1396
Db 1169 GTGCTCTCCAGACAGCGCTCTCAGTTCCTGCTGCTTGGCTTCTGCTTCTGCTTCTGCTTCT 1228
Qy 1397 ATCCTTACCCAAAGATTCAGACAAATCATGCTTGAAGCAAAACGCGGCTAACAGGA 1456
Db 1229 ATCCATACCTTAAGATTCGAGCAAGCAGACATGCTAGCTAAGCAAAACAGGACTCACTCGTA 1288
Qy 1457 GCCAGGTGCTTAAGTGTCTAATATGCTCGAGTTCGATTTATGGAAGCCATGGTAGAG 1516
Db 1389 GCCAGGTGCTCAACTGGTTTATAACCGGAGGTTCCGTTATGGAACCAATGGTGGAG 1348
Qy 1517 AAATGTAATTCGGAAGAGTGAAGAAATCAAGAAACAAACA 1555
Db 1349 AGATGTACATGGAGGAATGAAGAGCAGGCAAGAAACA 1387
```

RESULT 8

US-09-938-842A-1337
; Sequence 1337, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1337
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1337

Query Match 11.9%; Score 324.6; DB 11; Length 2043;

Best Local Similarity 65.2%; Pred. No. 5.1e-65;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

```
Qy 758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAGAGCTTCTTGATGAAGTTGTTAATATTG 817
Db 569 TTGTAGCTCCAACTACTTGAAGCGCACACAGAGCTTCTTGACGAAGTAGTCAACGCTG 628
Qy 818 TTGGAAGAGCTCAAGAGGAGATGATCAAAAGAGGATTAATCAATGAATAAAGAAATCAA 877
Db 629 ATTCCGATGACATGACGCTAATCCCACTATTCTCATCGAAAAGGGTAGTTGCGGAA 688
Qy 878 TGCCTTTGGCTAGTGTCAACACTAATATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 937
Db 689 ATGATAAACTGTCGAGAAATCATCGCGCGCTGAGGAGAGGTTCCGGTGGCGGAG 748
Qy 938 AGAAAAATG-----AAGTTGCTGTTGAGCTTCAAACTGCTCAAAAGCAAGAACTTCAAA 991
Db 749 CAGAAGCAGCGGGAACGTCGCTGAGCTTAGGCACGCGCAGAGAGACAAGAAATACAGA 808
Qy 992 TGAAGGAGGCAAGCTTCTTGCCTTGAAGAGGTGAGAGCAAGAGGTACAGACAGTACC 1051
Db 809 TGAAGGAGGCAAACTTAGTAACTATGCTTCAAGAGGTGAGAGAGATATAGACAGTACC 868
Qy 1052 ATCAACAAATGCAAAATATGTTATATCATTTGAGCAAGTAGCAGGAATGGATCAGCCA 1111
Db 869 ACCAGCAGATGCAGATGCTGATCTCTTCTGTCGAGCAAGCGCAGGAGTAGGATCAGCGA 928
Qy 1112 AATCATACACTCAATTTAGCTTTTGCATGCAATTTTGAAGCAATTTCAGATGCTTAAGAGATG 1171
Db 929 AGTCATACAGCTCGCTAGCATTTGAATCCATATCAAGACAGTTCCTGCTTGAAGAGG 988
Qy 1172 CAATTGCTGAGCAAGTAAAGCGCAGCAGAGAGTCTTAGGTGAAGAGGAGGCTTG--- 1227
Db 989 CGATCGCTGCTAGATTAAGCGGCAACAGAGTCTTGGGAGGAGATTCAGTGTCTG 1048
Qy 1228 --GGAGGGAATAACGAGGCTCAAGACTCAAAATTTTGGACCATCATCTAAGGCAACAA 1285
Db 1049 GTGTTGGAGGTTTGAAGGCTCAGGCTCAAGTTCGTTGGACCACTTGGAGAGAGG 1108
Qy 1286 GCGGCTGCAACAGATTCAGACAAATCATGCTTGAAGCAAAACGCGGCTAACAGGA 1336
Db 1109 GAGCTCTTCAACAACTCGGAATGATTCACATCCTTCCAAATATGCTTGGAGACCTCAAC 1168
Qy 1337 GAGGTTTACCTGAAAGAGTGTCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1396
Db 1169 GTGCTCTCCAGACAGCGCTCTCAGTTCCTGCTGCTTGGCTTCTGCTTCTGCTTCTGCTTCT 1228
Qy 1397 ATCCTTACCCAAAGATTCAGACAAATCATGCTTGAAGCAAAACGCGGCTAACAGGA 1456
Db 1229 ATCCATACCTTAAGATTCGAGCAAGCAGACATGCTAGCTAAGCAAAACAGGACTCACTCGTA 1288
Qy 1457 GCCAGGTGCTTAAGTGTCTAATATGCTCGAGTTCGATTTATGGAAGCCATGGTAGAG 1516
Db 1389 GCCAGGTGCTCAACTGGTTTATAACCGGAGGTTCCGTTATGGAACCAATGGTGGAG 1348
Qy 1517 AAATGTAATTCGGAAGAGTGAAGAAATCAAGAAACAAACA 1555
Db 1349 AGATGTACATGGAGGAATGAAGAGCAGGCAAGAAACA 1387
```



```
Db 807 ATCCGATGATGAACGCTAAATCCCACTATTCTCATCGAAAGGGTAGTTGCGGAA 866
Qy 878 TGCCTTTGGCTAGTAGTGTCAACACTAATAGTCTTGGTGGTGGTGAAGTAGCAGCAGGC 937
Db 867 ATGATAAACCTGTGCGAGATCATCGCGCGCTGGAGGAGAGGTTCCGGTGGCGAG 926
Qy 938 AGAAAAATG-----AAGTTGCTGTGAGCTTACAACTGCTCAAAAGACAAGAACTTCAAA 991
Db 927 CAGAGCAGCGCGGAAACGTCGCTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 986
Qy 992 TGAAGAAAGCCAACTCTTCCCATGCTTGAAGAGTGGAGCAAGTACAGACAGTACC 1051
Db 987 TGAAGAAAGCAAACTTAGTAACATGCTTCAAGAGTGGAGCAGAGATATAGACAGTACC 1046
Qy 1052 ATCAACAAATGCAAAATATGTTATCATTTGAGCAAGTAGCAGGAATTTGGATCAGCCA 1111
Db 1047 ACCAGCAGATGATGATGATCTCTTGGTTCGAGCAAGCGGAGGATAGATCAGCGA 1106
Qy 1112 AATCATACACTCAATATGCTTTGATGCAATTTTCAAGCAATTCAGATGCTTAAAGGATG 1171
Db 1107 AGTCATACAGCTGCTAGCATTTGAAACCATATCAAGACAGTTCCGTTGCTTGAAGAGG 1166
Qy 1172 CAATTGCTGAGCAAGTAAAGCGCAGCAGCAAGATTTAGTGAAGAGAGGCTTG----- 1227
Db 1167 CGATCGCTGGTTCAGATAAAGCGCGCAACAGAGCTTTGGGAGGAAAGATTCAGTGTCTG 1226
Qy 1228 --GGAGGAAATCGAAGGCTCAAGACTCAAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
Db 1227 GTGTTGGAGGTTTGAAGGCTGAGGCTCAAGTTCTGGACCACTTGTGAGACAGCAAA 1286
Qy 1286 GCGCGCTCAACAGATAGGAATGATCAAC-----CAATGCTTGGAGACCCCAAA 1336
Db 1287 GAGCTCTTCAACACTGCGATGATTCACATCCTTCCAAATATGCTTTGGAGACCTCAAC 1346
Qy 1337 GAGTTTACCTGAAGAGCTGTCTGTCTCTTCTGCTTCTGCTTCTGCTTCTGAGCATTTCTTC 1396
Db 1347 GTGCTCTCCAGAAACGAGCGCTCTCAGTTCTCCGCTGCTTGGCTTCTGCTTCTGCTTCT 1406
Qy 1397 ATCTTACCCAAAGGATTCAGACAAATCATGCTTGTGCTAGCAAAACGGGCTAACAGGA 1456
Db 1407 ATCCATACCTAAGATTCGACAAAGCAGATGCTAGCTAGCAAAACAGGACTCACTCGTA 1466
Qy 1457 GCCAGGTGCTAACTGCTTCAATATGCTCGAGTTCGATTTATGGAAGCAATGATGAAG 1516
Db 1467 GCCAGGTGCTGAACCTGTTTATAAACGCGAGGTTCCGTTATGGAACCAATGCTGAGG 1526
Qy 1517 AATGTACTTGGAAAGAGTGAAGATCAAGAACAAACA 1555
Db 1527 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAGAAACA 1565
```

RESULT 11

```
US-10-374-780A-2671
; Sequence 2671, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
```

```
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2671
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1589
US-10-374-780A-2671
```

```
Query Match 11.9%; Score 324.6; DB 17; Length 2385;
Best Local Similarity 65.2%; Pred. No. 5.6e-65;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

Qy 758 TTTTAGGCTCTAAGTATCTGAAAGCTGCAAGAGCTTCTTGTATGAAGTTGTTTAATATTG 817
Db 747 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 806
Qy 818 TTGAAAAAGCATCAAGGAGATGATCAAAAGAGGATAATTCATCAATGAATAAATCAA 877
Db 807 ATTCCGATGACATGAACGCTAAATCCCACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 866
Qy 878 TGCTTTGGCTAGTATGTCAACACTAATAGTTCTCGTGTGTGTAAGTAGCAGCAGGC 937
Db 867 ATGATAAACCTGTGCGAGATCATCGCGCGCTCGAGGAGAGGTTCCGGTGGCGAG 926
Qy 938 AGAAAAATG-----AAGTTGCTGTGAGCTTACAACTGCTCAAGACAAGAACTTCAAA 991
Db 927 CAGAGCAGCGCGGAAACGTCGCTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 986
Qy 992 TGAAGAAAGCCAACTTCTTGCATGCTTGAAGAGTGGAGCAAGAGTACAGACAGTACC 1051
Db 987 TGAAGAAAGCAAACTTAGTAACATGCTTTCATGAGTGGAGCAGAGATATAGACAGTACC 1046
Qy 1052 ATCAACAAATGCAAAATAATTTGATTTATCATTTGAGCAAGTAGCAGGAATTTGGATCAGCCA 1111
Db 1047 ACCAGCAGATGAGATGCTGATCTCTTGGTTCGAGCAAGCGGAGGATAGGATCAGCGA 1106
Qy 1112 AATCATACACTCAATTTAGCTTTGATGCAATTTTGAAGCAATTCAGATGCTTAAAGGATG 1171
Db 1107 AGTCATACAGCTGCTAGCATTTGAAACCATATCAAGACAGTTCCGTTGCTTGAAGAGG 1166
Qy 1172 CAATTGCTGAGCAAGTAAAGGCGCAGCAGAGGTTTAGTGAAGAGAGGCTTG----- 1227
Db 1167 CGATCGCTGCTCAGATAAAGCGCGCAACAGAGCTTCTGGGAGGAAAGATTCAGTGTCTG 1226
Qy 1228 --GGAGGAAATCGAAGGCTCAAGACTCAAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
Db 1227 GTGTTGGAGGTTTGAAGGCTGAGGCTCAAGTTCTGGACCACTTGTGAGACAGCAAA 1286
Qy 1286 GCGCGCTCAACAGATAGGAATGATCAAC-----CAATGCTTGGAGACCCCAAA 1336
```

Db 1287 GAGCTCTTCAACAACCTGGGAATGATTCACAATCCTTTCCAATATGCTTGGAGACCTCAAC 1346
QY 1337 GAGGTTTACCTGAAGAGCTGCTCTGTGCTCTGCTGCTTTCGAGCATTTTCTTC 1396
Db 1347 GTGGTCTCCAGAACGAGCGCTCTAGTTCCTCGTCTGGCTTTCGAACTTTCTTC 1406
QY 1397 ATCCTTTACCCAAAGGATTCAGACAAATCATGCTTCTAAGCAAAACGGGGCTTAACAAGGA 1456
Db 1407 ATCCATACCCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAAACAGGACTCACTCGTA 1466
QY 1457 GCAGGTGCTACTGCTGCTCAATAAATGCTCGAGTTCGATTTGGAAGCCAAATGGTAGAAG 1516
Db 1467 GCCAGGTGTCGAATCGGTTTATAAACCGGAGAGTTCGGTTATGGAACCAATGGTGGAGG 1526
QY 1517 AAATGTACTTGGAAAGTGAAGCAATCAAGAACAAAAA 1555
Db 1527 AGATGTACATGGAGGAATAGAGGACGAGCAAGAACA 1565

RESULT 12
US-10-225-066A-335
; Sequence 335, Application US/10225066A
; Publication No. US20050160493A9
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 335
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-335

Query Match 11.9%; Score 324.6; DB 22; Length 2385;
Best Local Similarity 65.2%; Pred. No. 5.6e-65;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;
QY 758 TTTTAGCTCTAGTATCTCAAGAGCTGCACAAGAGCTTCTTGATGAGTTGTTAATATG 817
Db 747 TTGTAGCTCAAGTACTTGAAGGACGACACAAGAGCTTCTTGACGAGTAGTCAACGCTG 806
QY 818 TTGMAAAGCATCAAGAGGAGATGATCAAAAGAGGATAATTCAATCAATAAAGAAATCAA 877
Db 807 ATTCCGATGACATGACGCTAAATCCCACTATTCTCATCGAAAAGGGTAGTTGCGGA 866
QY 878 TGCTTTGGCTAGTGTGTCAACACATAATATGTTCTGGTGTGTGTGAAAGTAGCAGGCG 937

Db 867 ATGATAAAACCTGTCCGAGAAATCATCGCGCGCGCTGGAGGAGAAAGGTTCCGGTGGCGGAG 926
QY 938 AGAAAAATG-----AAGTTGCTGTGAGCTTTACAACCTGCTCAAAAGACAAGAACTTCAAA 991
Db 927 CAGAACGAGCCGGGAAAACGTCCGGTGGAGCTAGGACACGGCAGAGAGACAAAGAAATACAGA 986
QY 992 TGAATAAAAGCCAAAGCTTCTTGCCATGCTTTGAAGAGGTGGAGCAAAAGGTCAGACAGTACC 1051
Db 987 TGAAGAAAGCAAAACCTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATACAGAGTACC 1046
QY 1052 ATCACCAATGCAAAATATTTATTCATTTGACCAAGTAGCAGAAATTTGATCAGCCA 1111
Db 1047 ACCAGCAGATCAGATGGTGATCTCTTCGTCGACCAAGCGCAGGGATAGGATCAGCGA 1106
QY 1112 AATCATACACTCAATTTAGCTTTGTCATGCAATTTTCAAGCAATTCAGATGCTCTAAAGGATG 1171
Db 1107 AGTCATACAGTCGCTAGCATTTGAAACCATATCAAGACAGTTCCTGCTGCTTTGAAGAGG 1166
QY 1172 CAATTGCTGAGCAAGTAAAGCGCAGCAGCAAGAGTTTAGGTGAAGAGGAGCTTG- --- 1227
Db 1167 CGATCGCTGTCAGATAAAAGCGGCCAACAGAGTCTTTGGGAGGAAGATTCAAGTGTCTG 1226
QY 1228 --GAGGGGAAATCGAAGCTCAAGACTCAAAATTTGTGGACCATCATCTAAGGCAACAC 1285
Db 1227 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCAACCACTTGACACAGCAA 1286
QY 1286 GCGCGCTGCAACAGATAGGAATGATGCAAC-----CAAATGCTTGGAGACCCCAAA 1336
Db 1287 GAGCTCTTCAACACTGGGAATGATTCACATCTCTCCATATGCTTGGAGACCTCAAC 1346
QY 1337 GAGGTTTACCTGAAAAGAGCTGCTCTGCTCTTCTGCTGCTTTGCGAGCAATTTCTTC 1396
Db 1347 GTGGTCTCCAGAACGAGCGCTCTCAGTTCTCCGTCTTGCTCTTCGAACACTTTCTTC 1406
QY 1397 ATCCTTACCCAAAGGATTCAGACAAATCATGCTTCTAAGCAAAAGGGGCTTAACAAGA 1456
Db 1407 ATCCATACCTTAAGGATTCGGAACAGCACATCTAGCTAAGCAAAACAGGACTCACTCGTA 1466
QY 1457 GCAGGTGCTAACTGGTTTCATAAATGCTCGAGTTTCGATTATGGAAGCCAAATGGTAGAAG 1516
Db 1467 GCAGGTGCTGAACCTGGTTTATAAAGCGGAGAGTTGCGTTATGGAACCAATGGTGGAGG 1526
QY 1517 AAATGTACTTGAAGAAAGTGAAGAAATCAAGAACAAAAA 1555
Db 1527 AGATGTACATGAGGAATGAAGGAGCAGGCAAGAACA 1565

RESULT 13
US-10-425-114-9882
; Sequence 9882, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9882
; LENGTH: 1716
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700875318_FLI
US-10-425-114-9882

Query Match 11.5%; Score 314; DB 18; Length 1716;

```

RESULT 14
US-10-424-599-7028
; Sequence 7028, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 7028
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106354C.1
US-10-424-599-7028

Query Match      11.1%; Score 304; DB 18; Length 582;
Best Local Similarity 76.4%; Pred. No. 1.7e-60;
Matches 373; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy    957 TGAGCTTACAACGTGCTCAAGAGACAAGAAGACTTCAAATGAAAAAGGCCAACGTTCTTGCCAT 1016
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    95 TGAATTGTTCAACTACGGAGAGACAAGAATAATTCAGATGAGAGACCACCAATCATCGACAT 154
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy    1017 GCTTGAAGAGGTTGGAGCAAAGGTCAGACAGTAGTACCATCACCAATCAAATATTGTATT 1076

```

```

? APPLICANT: Cao, Tongwei
? APPLICANT: Wu, Wei
? APPLICANT: Boukharov, Andrey A.
? APPLICANT: Barbazuk, Brad
? APPLICANT: Li, Ping
? TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
? TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
? FILE REFERENCE: 38-21(53221)B
? CURRENT APPLICATION NUMBER: US/10/437,963
? CURRENT FILING DATE: 2003-05-14
? NUMBER OF SEQ ID NOS: 204966
? SEQ ID NO 41007
? LENGTH: 3028
? TYPE: DNA
? ORGANISM: Oryza sativa
? FEATURE:
? OTHER INFORMATION: Clone ID: PAT_MRT4530_44394C.1
US-10-437-963-41007

Query Match          10.8%;   Score 295.6;   DB 19;   Length 3028;
Best Local Similarity 60.3%;   Pred No. 3.9e-58;
Matches 528;   Conservative 0;   Mismatches 339;   Indels 9;   Gaps 2;

Qy      765   CTCCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAAATTATGTTGGAAA 824
      |||
Db      1135  CTCGAAGGTATCTCAAGGCAGCACAGAAGCTGCTTGATGAGGTCGTGAGTGTGTTGGAAGAG 1194

Qy      825   AAGCATCAAGGAGGATGATCAAAAGAGGATAATTCAATGAATAAAGATCAATGCCTTT 884
      |||
Db      1195  CATTAACAAGAAGCTCAGAAGAAGAGAGTTCAATCAGGAAAAGCAGATGGCAAA----- 1249

Qy      885   GGCTAGTGATCTCAACATAATAGTCTTGTTGGTGGTGAAGTAGCAGCAGGCAGAAAAA 944
      |||
Db      1250  -GAGACTGATGGAGGGCCAAAAGAGTGAAGGGTGTATCTCTCTAATCCACAAGAATCTGCGCG 1308

```

Qy	945	TGAAGTTGCTGTGAGCTTACAACTGCTCAAGACAAGAACTTCAAAATGAAAAAGCCAA	1004
Db	1309	TAATGACGACCTGAGCTTCCACTGCTGGAAGCAAGAGCTTCAGAACAAATGSCAAA	1368
Qy	1005	GCTTCTTGCCATGCTTGAAGAGTGAGCAAAAGGTACAGACAGTACCATCACCAAATGCA	1064
Db	1369	ACTGATGGCAATGTTGGATGAGGTGACCGGAAATACAAAGCAATTAATCCACCAAATGCA	1428
Qy	1065	ATAAATTGATATATATTTGAGCAAGTACAGCAATTTGGATCAGCCAAATCATACACTCA	1124
Db	1429	AACCGTGGTTTCATCGCTTTGATGTGGTAGCTGGACCAGGATCTCGAAGCCTTACACAGC	1488
Qy	1125	ATTAGCTTTGTCATGCAATTTGGAAGCAATTCAGATGCTTAAAGGATGCAATTCCTGAGCA	1184
Db	1489	AGTTGCTCTTCAGACAAATCTCAGCGCACTTCGGTGCCTGAAGGATGCTATCAATGATCA	1548
Qy	1185	AGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGCTTGGGAGGGAAAAATCGAAGG	1244
Db	1549	GATCAATGTTATCAGGAAAAAGCTTGGCGAGGAAGAGAACTCATCTGGCAAGGAGGGAAA	1608
Qy	1245	CTCAA---GACTCAAAATTTGTGACCATCATCTAAGGCAACACGCGCTGCAACAGAT	1301
Db	1609	ATTAAACCCGCTCCGTACATTGATCAGCAGCTTAAGGCAACACGCTGCTTCCACAGTA	1668
Qy	1302	AGGAATGATGCAACCAAAATGCTTGGAGACCCCAAGAGGTTTACCTGAAAAGAGCTGCTC	1361
Db	1669	CGGTATGATTTCCACAAAAACGCTGGAGACCAACAGAGGGAGCTGCTGAAAACTCTGTTAC	1728
Qy	1362	TGTCCTTGCTGTGCTTTTGAGCAATTTTCTCATCTTACCCAAAGGATTCAGACAA	1421
Db	1729	AATCTTCGCGCTTGGCTGTTTGAACACTTCTCCATCCGTACCCAAAAGATTCCGAGAA	1788
Qy	1422	AATCATGCTTGCTAAGCAACGGGGCTAACAGAGGCCAGGTGCTTAATGGTTTCATAA	1481
Db	1789	GTTAATGCTTGTAGGAGACTGGCTGACTAGGAGTCAGATTTCGAATTTGGTTTCATAA	1848
Qy	1482	TGCTCGAGTTCCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAAGAGTGAAGAA	1541
Db	1849	TGCCCCGTGTCGGCTTTGGAAACCAATGATTGAAGACATGTACAAGAGAGAGATTGGTGA	1908
Qy	1542	TCAAGAACAAAAACAGTACTTAATCTTCAGAGAGATAACAAAAACAAAGAGACCAATATAAG	1601
Db	1909	TTTAGAGCAAGACTCCAACTCTTCTCTGACAAATGCACCAAGAAAGTAAGGATAAAATGGC	1968
Qy	1602	TGCTCCAAATGAGAGAAACATCCATTTACTAG	1637
Db	1969	ATCTTCAGAGACAGAGGATCTCAAAAACCTCCAG	2004

Search completed: July 29, 2005, 14:48:36
Job time : 1683 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 07:12:43 ; Search time 8585 Seconds
(without alignments)
12126.474 Million cell updates/sec

Title: US-10-624-201a-1
Perfect score: 2735
Sequence: 1 catcgagagataaaatata.....gcaaaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	776.4	28.4	786	5 BQ509373	BQ509373 EST616788
c 2	722.2	26.4	743	5 BQ510344	BQ510344 EST617759
c 3	718.6	26.3	756	4 BG593861	BG593861 EST492539
c 4	666.4	24.4	699	5 BQ119770	BQ119770 EST605346
c 5	638.6	23.3	685	2 AW034617	AW034617 EST278301
c 6	608	22.2	708	4 B1920940	B1920940 EST540875
c 7	594.8	21.7	727	4 B1176426	B1176426 EST517371
c 8	581.6	21.3	644	2 AW930221	AW930221 EST340678
c 9	565	20.7	630	4 BQ129304	BQ129304 EST474950
c 10	564.2	20.6	789	4 B1921008	B1921008 EST540943
c 11	531	19.4	610	7 CQ909070	CQ909070 BJ02039F0
c 12	517.8	18.9	553	1 A1773253	A1773253 EST254353
c 13	502.4	18.4	721	2 BQ435130	BQ435130 EST406208
c 14	476	17.4	964	7 CK262102	CK262102 EST708180
c 15	467.8	17.1	495	2 BQ460264	BQ460264 EST411755
c 16	467.2	17.1	480	5 BQ509372	BQ509372 EST616787
c 17	446.2	16.3	623	2 AW441343	AW441343 EST310739
c 18	438	16.0	747	7 CV472924	CV472924 21081.1 D
c 19	427.8	15.6	479	2 BQ433793	BQ433793 EST404871
c 20	424.8	15.5	452	2 AW041164	AW041164 EST284028
c 21	419	15.3	752	4 BM535737	BM535737 EST588759
c 22	417.8	15.3	568	2 AW441354	AW441354 EST310750
c 23	417.6	15.3	440	2 AW093664	AW093664 EST286844
c 24	417.4	15.3	452	2 BF188184	BF188184 EST444471

25	403.8	14.8	576	1	AI894617	AI894617 EST264060
26	393.8	14.4	667	1	AI485358	AI485358 EST243679
27	392	14.3	740	7	CO111273	CO111273 GR_EB004
28	391.4	14.3	554	4	AI897675	AI897675 EST267118
29	391.4	14.3	575	4	BI926891	BI926891 EST546780
30	389.6	14.2	494	2	AW094308	AW094308 EST287488
31	389.4	14.2	530	1	AI487525	AI487525 EST245847
32	388.6	14.2	523	1	AI898911	AI898911 EST268354
33	383.8	14.0	669	5	BQ148285	BQ148285 NF065F11F
34	383	14.0	783	2	AW688195	AW688195 NF004R06S
35	382.4	14.0	415	7	CV494999	CV494999 40158.1 C
36	377.2	13.8	645	4	EG887850	EG887850 EST513701
37	367.2	13.4	850	6	CB292855	CB292855 UCRCS01.0
38	366.8	13.4	1064	7	CV477853	CV477853 57798.1 D
39	363.6	13.3	844	7	CO101205	CO101205 GR_EB002
40	361	13.2	636	4	EG887971	EG887971 EST513822
41	360.8	13.2	737	2	BE435482	BE435482 EST406560
42	356	13.0	529	4	BI930473	BI930473 EST550362
43	354.6	13.0	537	5	BM954132	BM954132 sam70e12.
44	352.4	12.9	636	2	AW442048	AW442048 EST311444
45	350.2	12.8	550	2	BE341111	BE341111 EST345206

ALIGNMENTS

RESULT 1
LOCUS BQ509373/c 786 bp mRNA linear EST 07-MAR-2003
DEFINITION EST616788 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMHC24 3' end, mRNA sequence.
ACCESSION BQ509373
VERSION BQ509373.2 GI:21925113
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum.
AUTHORS Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C., Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and Karamycheva,S.A.
TITLE Generation of a set of potato cDNA clones for microarray analyses
JOURNAL Unpublished (2002)
COMMENT On Jun 10, 2002 this sequence version replaced gi:21368242.
Other_ESTs: EST616787
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL: http://genome.arizona.edu/orders/
Seq primer: T7.
Location/Qualifiers
1. .786
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="kennebec or Binjite"
/db_xref="taxon:4113"
/clone="STMHC24"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: xhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."

ORIGIN

Query Match		28.4%;	Score 776.4;	DB 5;	Length 786;
Best Local Similarity		99.2%;	Pred. No. 4.5e-172;		
Matches 780;		Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	1881	AGCTAGAGAAATCATCAAAATAAAGGGTTTACTAATCCCTTTAAATGGCAGCATACCGATGGG	1940		
DB	786	AGCTAGAGAAATCATCAAAATAAAGGGTTTACTATCCCTTTAAATGGCAGCATACCGATGGA	727		
QY	1941	AGATTTTGGAGGTTTGATCTCTCATGATCAACAAATGACCGCGAATTTTCATGGAATAA	2000		
DB	726	GATTTTGGAGGTTTGATCTCTCATGATCAACAAATGACCGCGAATTTTCATGGAATAA	667		
QY	2001	TGGTGCTCTCTTACCTTTAGGACTTCTCTCTCTGAAACCTAGCCATGCCAGTGAGCCA	2060		
DB	666	TGGTGCTCTCTTACCTTTAGGACTTCTCTCTCTGAAACCTAGCCATGCCAGTGAGCCA	607		
QY	2061	ACAAAATTTACCTTTCTAATGACTTGGGAAGTAGTCTGAAATGGGAGTCATTACAATAG	2120		
DB	606	ACAAAATTTACCTTTCTAATGACTTGGGAAGTAGTCTGAAATGGGAGTCATTACAATAG	547		
QY	2121	AATGGGATATGAACAATTTGATTTTCAGAGTGGGAATGACCGATTTTCCGACTCAACTATT	2180		
DB	546	AATGGGATATGAACAATTTGATTTTCAGAGTGGGAATGACCGATTTTCCGACTCAACTATT	487		
QY	2181	ACCAGATTTTGTGTACAGGTAATCTAGGAACATCAATACCAAGAAAGTCTCGTATTGTATAGC	2240		
DB	486	ACCAGATTTTGTGTACAGGTAATCTAGGAACATCAATACCAAGAAAGTCTCGTATTGTATAGC	427		
QY	2241	TGAAAAGATAAAAGGAAGTTAGGGATACCTCTTATTTGTGTAGGGCTTCTTGSCCCTCAAGT	2300		
DB	426	TGAAAAGATAAAAGGAAGTTAGGGATACCTCTTATTTGTGTAGGGCTTCTTGSCCCTCAAGT	367		
QY	2301	CGAGAGACCAATTTGATCAACCTATCATAGGAGAAAGAAAGTGAGACTAAATTTAAAG	2360		
DB	366	CGAGAGACCAATTTGATCAACCTATCATAGGAGAAAGAAAGTGAGACTAAATTTAAAG	307		
QY	2361	TAACAAAATTTTAAAGCACACTTTCTAGTATATATATATCTCTTTTATAGTATAGAAA	2420		
DB	306	TAACAAAATTTTAAAGCACACTTTCTAGTATATATATATCTCTTTTATAGTATAGAAA	247		
QY	2421	AGAAGAGATTTTGTCTTTAGTGTATAGATAGAGTCTACTTAGTAGTATAGTTTACTTA	2480		
DB	246	AGAAGAGATTTTGTCTTTAGTGTATAGATAGAGTCTACTTAGTAGTATAGTTTACTTA	187		
QY	2481	GTTCTCTGAGAAATTTGATCAACTAGTAGTATTTTCTTTTGGGTTGGCTTCGAGT	2540		
DB	186	GTTCTCTGAGAAATTTGATCAACTAGTAGTATTTTCTTTTGGGTTGGCTTCGAGT	127		
QY	2541	ACTATTTTAAAGTTATTGGAACTAGCTATAGTAAATGTTGTAAGTTGTGATATTCTTC	2600		
DB	126	ACTATTTTAAAGTTATTGGAACTAGCTATAGTAAATGTTGTAAGTTGTGATATTCTTC	67		
QY	2601	TCTCAATTTGCAATTAATTTGAAATATTTTGTACCTACTAGCTAGTCTCTAAATTAAGTT	2660		
DB	66	TCTCAATTTGCAATTAATTTGAAATATTTTGTACCTACTAGCTAGTCTCTAAATTAAGTT	7		
QY	2661	TCCATT 2666			
DB	6	TCCATT 1			
RESULT 2					
BQ510344/c					
LOCUS					
DEFINITION					
BQ510344 Generation of a set of potato cDNA clones for microarray					
analyses mixed potato tissues Solanum tuberosum cDNA clone STMHJ58					
3' end, mRNA sequence.					
ACCESSION					
BQ510344					
VERSION					
BQ510344.2					
KEYWORDS					
EST.					
Solanum tuberosum (potato)					
SOURCE					
Solanum tuberosum					
ORGANISM					
Solanum tuberosum					
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.					
1 (bases 1 to 743)					
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and Karamycheva, S.A.					
Generation of a set of potato cDNA clones for microarray analyses					
Unpublished (2002)					
On Jun 10, 2002 this sequence version replaced gi:21369213.					
Other ESTs: EST617758					
Contact: Robin Buell					
The Institute for Genomic Research					
9712 Medical Center Dr, Rockville, MD 20850, USA					
Email: potato-array@tigr.org					
This clone can be obtained from the University of Arizona Genomics					
Institute. Orders can be made through URL:					
http://genome.arizona.edu/orders/					
Seq primer: T7.					
Location/Qualifiers					
1..743					
/organism="Solanum tuberosum"					
/mol_type="mRNA"					
/cultivar="Kennebec or Binjite"					
/db_xref="taxon:4113"					
/clone="STMHJ58"					
/tissue_type="mixed tissues"					
/lab_host="SOLR"					
/clone_lib="Generation of a set of potato cDNA clones for					
microarray analyses mixed potato tissues"					
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:					
XhoI; supplier: Combination of untreated and Phytophthora					
infestans-treated libraries of stolons, leaves, leaflets,					
axillary buds of stem explants, petioles, germinating					
eyes, tubers, or roots."					
ORIGIN					
Query Match 26.4%; Score 722.2; DB 5; Length 743;					
Best Local Similarity 99.5%; Pred. No. 2.8e-159;					
Matches 735; Conservative 0; Mismatches 3; Indels 1; Gaps 1;					
QY	1811	AAAAGCAAAGAAATGACATGCACAAAGTTTCTCCCAAGTAGTATTTCTTCATCTGTGACA	1870		
DB	743	AAAAGCAAAGAAATGACATGCACAAAGTTTCTCCCAAGTAGTATTTCTTCATCTGTGACA	684		
QY	1871	TGGAAGCCAAAGCTAGAGAAATCAATCAATAAAGGGTTTACTAATCCTTTAATGGCAGCAT	1930		
DB	683	TGGAAGCCAAAGCTAGAGAAATCAATCAATAAAGGGTTTACTAATCCTTTAATGGCAGCAT	624		
QY	1931	ACGCGATGGAGATTTTGGAGGTTTGATCCCTCATGATCAACAAATGACCGGAAATTTTC	1990		
DB	623	ACGCGATGGAGATTTTGGAGGTTTGATCCCTCATGATCAACAAATGACCGGAAATTTTC	564		
QY	1991	ATGGAATAATGGTGCTCTCTTACTTTAGGACTTCTCTCTCTGAAAACTAGCCATGC	2050		
DB	563	ATGGAATAATGGTGCTCTCTTACTTTAGGACTTCTCTCTCTGAAAACTAGCCATGC	504		
QY	2051	CAGTGAGCCAAACAAATTTTCTAATGACTTGGGAAGTAGTCTGAAATGGGGAGTC	2110		
DB	503	CAGTGAGCCAAACAAATTTTCTAATGACTTGGGAAGTAGTCTGAAATGGGGAGTC	444		
QY	2111	ATTACAATAGATGGGATATGAAACATTGATTTTTCAGAGTGGGAATAGCCGATTTCCGA	2170		
DB	443	ATTACAATAGATGGGATATGAAACATTGATTTTTCAGAGTGGGAATAGCCGATTTCCGA	384		
QY	2171	CTCAACTATTACCAGATTTTGTACAGGTAATCTAGGAAATGAAATGACAGAAAGTCTCG	2230		
DB	383	CTCAACTATTACCAGATTTTGTACAGGTAATCTAGGAAATGAAATGACAGAAAGTCTCG	324		
QY	2231	TATTGATAGCTGAAAGAGATAAAGGAGTTAGGATCTCTTATATTGTGAGGCTTC	2290		
DB	323	TATTGATAGCTGAAAGAGATAAAGGAGTTAGGATCTCTTATATTGTGAGGCTTC	264		

```
QY 2291 TGCCCCAAGTCGGAGGACCAAAATTTGATACAACTATCATAGAGAAAAGAACTGGAGAC 2350
Db 263 TGCCCCAAGTCGGAGGACCAAAATTTGATACAACTATCATAGAGAAAAGAACTGGAGAC 204
QY 2351 TAAATTAAAGTAACAAATTTTAAAGCACACTTCTAGTATATATATCTCTCTTTTAT 2410
Db 203 TAAATTAAAGTAACAAATTTTAAAGCACACTTCTAGTATATATATCTCTCTTTTAT 144
QY 2411 AGTATAGAAAAGAGAGATTTTGTGCTTTAGTGTATAGATAGAGTCTCTTAGTATAGGT 2470
Db 143 AGTATAGAAAAGAGAGATTTTGTGCTTTAGTGTATAGATAGAGTCTCTTAGTATAGGT 84
QY 2471 TATACCTCTAGTCTCTTGAGAGATTGATACAACTAGTAGTATTTTCTTTTGGGTT 2530
Db 83 TATACCTCTAGTCTCTTGAGAGATTGATACAACTAGTAGTA-TTTTCTTTCTTTGGGTT 25
QY 2531 GGCTTGGAGTACTATTTTA 2549
Db 24 GGCTTGGAGTACTATTTTA 6

RESULT 3
BG593861
LOCUS BG593861
DEFINITION BG593861 756 bp mRNA linear EST 07-MAR-2003
mRNA sequence.
ACCESSION BG593861
VERSION BG593861.1 GI:13612001
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
REFERENCE 1 (bases 1 to 756)
AUTHORS Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R
FEATURES
source
1..756
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTSP9"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/clone_lib="cSTS"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
```

ORIGIN

```
Query Match 26.3%; Score 718.6; DB 4; Length 756;
Best Local Similarity 98.1%; Pred. No. 2e-158;
Matches 741; Conservative 0; Mismatches 4; Indels 10; Gaps 1;
QY 41 AACTTCTCAAGCTTAGAGCTTACCACCGAGATAGACAGTTAGTTACA----- 91
Db 2 AACTTCTCAAGCTTAGAGCTTACCACCGAGATAGACAGTTAGTTACATGTGTAT 61
```

```
QY 92 -TGTACTGTTATAGATAAAAGGAGAAATCCGAAGAAAGAAATTTTTTTTGCAGATATG 150
Db 62 GTGTACTGTTATAGATAAAAGGAGAAATCCGAAGAAAGAAATTTTTTTTGCAGATATG 121
QY 151 TACTATCAAGGAACCTCGGATATATCTATATATCAAGCTGATCATCAACACGCTCATAT 210
Db 122 TACTATCAAGGAACCTCGGATATATCTATATATCAAGCTGATCATCAACACGCTCATAT 181
QY 211 CATGGAAATAGTAATATATATATATATATATATATATATATATATATATATATATAT 270
Db 182 CATGGAAATAGTAATATATATATATATATATATATATATATATATATATATATATAT 241
QY 271 ATGCAAGGCTACACTTCTGACACACAGCAGCAGCAGTACTTTTCTGGAATTCT 330
Db 242 ATGCAAGGCTACACTTCTGACACACAGCAGCAGCAGTACTTTTCTGGAATTCT 301
QY 331 TCACCAGCAGCAAGCAACGGCTTTGCCATGCGAATATACACACGGCGCTGCAACAG 390
Db 302 TCACCAGCAGCAAGCAACGGCTTTGCCATGCGAATATACACACGGCGCTGCAACAG 361
QY 391 CAGCACTTTGTGCGTGTGCTCTTCCGCGAGTAAGTTTGACAGATCAGATCAATCATCAT 450
Db 362 CAGCACTTTGTGCGTGTGCTCTTCCGCGAGTAAGTTTGACAGATCAGATCAATCATCAT 421
QY 451 GGACTTTTACAGCGCATGTGGAACCAACCAAGATCAATCTCAGCAGGTGATAGTACCATCG 510
Db 422 GGACTTTTACAGCGCATGTGGAACCAACCAAGATCAATCTCAGCAGGTGATAGTACCATCG 481
QY 511 TCAGCGGGGTTTCTGCCAGCTCATGTGGGGGATCACCACGAGTCTGGCTCTCAATTG 570
Db 482 TCAGCGGGGTTTCTGCCAGCTCATGTGGGGGATCACCACGAGTCTGGCTCTCAATTG 541
QY 571 GCGTTTTCAGAGCGCGATTCGACACCAACACCGCAGCAGCAACAAAGAGGCGGT 630
Db 542 GCGTTTTCAGAGCGCGATTCGACACCAACACCGCAGCAGCAACAAAGAGGCGGT 601
QY 631 CTATCTCTAAGCCTTTCTCTCAGCTACCAACAGCAAAATTAGTTTCAATAATATTTC 690
Db 602 CTATCTCTAAGCCTTTCTCTCAGCTACCAACAGCAAAATTAGTTTCAATAATATTTC 661
QY 691 TCTCTATCACAAGGCAATATATGTTACTATNTAGGGGAACATTAGATGGAAGTTCTAGC 750
Db 662 TCTCTATCACAAGGCAATATATGTTACTATNTAGGGGAACATTAGATGGAAGTTCTAGC 721
QY 751 AACATGGTTTATAGGCTCTAAGTATCTGAAAGCTGC 785
Db 722 AACATGGTTTATAGGCTCTAAGTATCTGAAAGCTGC 756
```

RESULT 4

BG119770

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BQ119770 699 bp mRNA linear EST 07-MAR-2003
EST605346 mixed potato tissues Solanum tuberosum cDNA clone STWEM54
5' end, mRNA sequence.

BQ119770
BQ119770.2 GI:21919664

EST.

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 699)

Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karamycheva,S.A.

Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)

On Apr 17, 2002 this sequence version replaced gi:20171732.

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Db 301 AAGACTCAGACAAAATCATGCTTGTCTAAGCAACGGGGCTTAAACAGAGCCAGTCTCT 360
Qy 1468 AACTGGTTCTAAATGCTCGAGTTTCGATTATGGAAGCAATGGTAGAAGAAATGACTTGT 1527
Db 361 AACTGGTTCTAAATGCTCGAGTTTCGATTATGGAAGCAATGGTTGAAGAAATGACTTGT 420
Qy 1528 GAAGAAGTGAAGATCAAGAACAAACAGTACTTAATCTTCAGGAGATTAACAAAACAAA 1587
Db 421 GAAGAAGTGAAGATCAAGAACAAACAGTACTTAATCTTCAGGAGATTAACAAAACAAA 480
Qy 1588 GAGACCAATATAAGTGTCCAAATGAAGAGAAACATCCAAATTTATTACTAGCAGCTTATTA 1647
Db 481 GAGACGAATATAGTGTCCAAATGAAGAGAAACAAATTTATTACTAGCAGCTTATTA 540
Qy 1648 CAAGATGGTATTACTACTACTCAGACGAAATTTCTACCTCAACTATTTCAACTTCCCT 1707
Db 541 CAAGATGG-----TACTACTCAGACGAAATTTCTACCTCAACTATTTCAACTTCCCT 594
Qy 1708 ACTGCAGGTGTTCACTTTCATCTCATGCTCACAATTTCTCTTCTTCTGTTTCATTCAACATG 1767
Db 595 ACTGCAGGTGTTCACTTTCATCTCATGCTCACAACCTTCTCTTCTTCTGTTTCATTCAACATG 654
Qy 1768 GATAATACTACTACTACTCTGTTGATCATATTG 1798
Db 655 GAGAACTACTACTACTCTGTTGATCATATTG 685

RESULT 6

BI920940 708 bp mRNA linear EST 10-MAR-2003
LOCUS EST540875 potato microtubers, in vitro-grown Solanum tuberosum cDNA
DEFINITION clone cSTE26N18 5' end, mRNA sequence.

ACCESSION BI920940

VERSION BI920940.1 GI:16216968

KEYWORDS EST.

SOURCE Solanum tuberosum (potato)

ORGANISM

Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 708)

AUTHORS van der Hoeven, R., Bezzerides, J., Bachem, C., Visser, R.,
Karamycheva, S.A., Tsai, J., Van Aken, S., Uterback, R., Chieningo, A.,
Bougr, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.
Generation of ESTs from in vitro grown microtubers (2001b)
Unpublished (2001)

CONTACT: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:

http://genome.arizona.edu/orders/

Seq primer: T3

Location/Qualifiers

1. 708

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Bintje"

/db_xref="taxon:4113"

/clone="cSTE26N18"

/tissue_type="axillary buds of stem explants; growing

sink-tubers"

/dev_stage="7, 8 and 10 days"

/lab_host="SOLR"

/clone_lib="potato microtubers, in vitro-grown"

/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; Tissue supplied by Christian Bachem and Richard
Visser (Department of Plant Breeding, Wageningen
University, The Netherlands); sequencing by The Institute
for Genomic Research. The cSTA libraries will attempt to
capture the induction and initiation/initial growth of the
tuber in an in vitro system as described in Bachem et al.

(Plant Journal, 1996). Small microtubers develop from
axillary buds attached to stem explants when placed on a
high sucrose medium (10%). Visible morphological changes
occur synchronously at day five in the axillary buds. The
first library, cSTA (1-20) consists of axillary buds
harvested on days 1-3. This targets those genes involved
in induction of the microtubers. The following libraries,
cSTA (21-40) and cSTA (41-60), capture genes involved in
tuber initiation and outgrowth. This library is noted as
P3 in Tanksley lab notebooks."

ORIGIN

Query Match 22.2%; Score 608; DB 4; Length 708;

Best Local Similarity 94.2%; Pred. No. 2.5e-132;

Matches 669; Conservative 0; Mismatches 25; Indels 16; Gaps 3;

Qy 50 AAGCTTAGAGAGCTACCCCGAAGATAGACAGCTAGTTACA-----TGTACTGT 99
Db 2 AAGTTTAGAGAGCTACCCCGAAGATAGACAGCTAGTTACA-----TGTACTGT 61
Qy 100 TATAGATAAAGGAGAAATCCGAAGAAAGAAATTTTTTTCAGATATGACTATCAA 159
Db 62 TATAGATAAAGGAGAAATCCGAAGAAAGAAATTTTTTTCAGATATGACTATCAA 121
Qy 160 GGAACCTCGGATAATCTAATATACAAGCTGATCATCAACACGTCATATCATGGAAT 219
Db 122 GGAACCTCGGATAA---TAATATACAAGCTGATCATCAACACATCATATCATGGAAT 178
Qy 220 AGTAATAATAATAATTCAGACACTTTATTCATGAACCTTAACCAATATATGCAAGGC 279
Db 179 AGTAATAATAATAATTCAGACACTTTATTCATGAACCTTAACCAATATATGCAAGGC 238
Qy 280 TACACTACTTCTGACACA---CAGCAGCAGCAGCAGTACTTTTCTGAAATTTCTCACCA 336
Db 239 TACACTACTTCTGACACACAGCAGCAGCAGCAGTACTTTTCTGAAATTTCTCACCA 298
Qy 337 CGAGCAAGCAACCGCTTTGCGATGCAATATACAACACGCGCGCTGCAACAGCAGCAC 396
Db 299 GCAGGAAGCAACCGCTTTGCGATGCAATATACAACACGCGCGCTGCAACAGCAGCAC 358
Qy 397 TTTGTGCGTGTGCTCTTCGCGCAGTAAGTTTCACGATCAGATCAATCATCATGACTT 456
Db 359 TTTGTGCGTGTGCTCTTCGCGCAGTAAGTTTCACGATCAGATCAATCATCATGACTT 418
Qy 457 TTACAGCGCATGTGGAACCAACCAAGATCAATCTCAGCAGGTGATAGTACCATCGTCGACG 516
Db 419 TTACAGCGCATGTGGAACCAACCAAGATCAATCTCAGCAGGTGATAGTACCATCGTCGACG 478
Qy 517 GGGGTTTTCGCCACGTCATGTGGCGGATCACACGAGCTTGGCGTCTCAATTGGCGTTT 576
Db 479 GGGGTTTTCGCCACGTCATGTGGCGGATTTACCACGAGCTTGGCGTCTCAATTGGCGTTT 538
Qy 577 CAGAGCGCGATTCGACACACACACACCGCAGCAGCAACAGCAAGCGCGTCTATCT 636
Db 539 CAGAGCGCGATTTGTTGTGCGAACCACCGCAGCAGCAACAGCAAGCGCGTCTATCT 598
Qy 637 CTAAGCGCTTTCTCCTCAGCTACAAACAGCAAAATTTAGTTTCAATAACCAATATTTTCATCTCA 696
Db 599 CTAAGCGCTTTCTCCTCAGCAACCAACAGCAAAATTTAGTTTCAATAACCAATATTTTCAGCCTCA 658
Qy 697 TCACCAAGGACAATAATGTTACTATTAGGGGAACATTAGATGGAAGTTC 746
Db 659 TCACCAAGGACAATAATGTTACTATTAGGAAGATTAGATGGAAGTTC 708

RESULT 7

BI176426

LOCUS

DEFINITION

EST517371 cSTE Solanum tuberosum cDNA clone cSTE1A17 5' sequence,

mRNA sequence.

ACCESSION BI176426

VERSION BI176426.1 GI:14642237

KEYWORDS EST.

BI176426 727 bp mRNA linear EST 07-MAR-2003
EST517371 cSTE Solanum tuberosum cDNA clone cSTE1A17 5' sequence,
mRNA sequence.

```

SOURCE
ORGANISM      Solanum tuberosum (potato)
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS       1 (bases 1 to 727)
               van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
               Chieningo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and
               Baker,B.
               Generation of ESTs from in vitro grown microtubers
               Unpublished (2001)
               Contact: Robin Buell
               The Institute for Genomic Research
               9712 Medical Center Dr, Rockville, MD 20850, USA
               Email: potato-array@igr.org
               This clone can be obtained from the University of Arizona Genomics
               Institute. Orders can be made through URL:
               http://genome.arizona.edu/orders/
               Seq primer: M13F-R.
FEATURES
source        Location/Qualifiers
               1..727
               /organism="Solanum tuberosum"
               /mol_type="mRNA"
               /cultivar="Bintje"
               /db_xref="taxon:4113"
               /clone="cSTEL17"
               /tissue_type="axillary buds of stem explants; growing
               sink-tubers"
               /dev_stage="7, 8 and 10 days"
               /lab_host="SOLR"
               /clone_lib="cSTE"
               /notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
               XhoI; Tissue supplied by Christian Bachem and Richard-
               Visser (Department of Plant Breeding, Wageningen
               University, The Netherlands). The cSTA libraries will
               attempt to capture the induction and initiation/initial
               growth of the tuber in an in vitro system as described in
               Bachem et al. (Plant Journal, 1996). Small microtubers
               develop from axillary buds attached to stem explants when
               placed on a high sucrose medium (10%). Visible
               morphological changes occur synchronously at day five in
               the axillary buds. The first library, cSTA (1-20) consists
               of axillary buds harvested on days 1-3. This targets
               those genes involved in induction of the microtubers. The
               following libraries, cSTA (21-40) and cSTA (41-60),
               capture genes involved in tuber initiation and outgrowth.
               This library is noted as P3 in tanksley lab notebooks."
ORIGIN
Query Match      21.7%; Score 594.8; DB 4; Length 727;
Best Local Similarity 94.9%; Pred. No. 3.3e-129;
Matches 695; Conservative 0; Mismatches 17; Indels 20; Gaps 7;

QY 17 TATAGATCAGCTGTGACAAGGCAACTTCTCAAAGCTTAGAGAGTACACCCCGAAGAT 76
DB 1 TATAGATCAGCTGTGACAAGGCAACTTCTCAAAGCTTAGAGAGTACACCCCGAAGAT 60
QY 77 AGACAGTGTAGTTACA-----TGTACTGTTATAGATAAAGAGAAATCCGAGAA 126
DB 61 AGACAGTGTAGTTACATGTGTATGTGTACTGTTATAGATAAAGAGAAATCCGAGAA 120
QY 127 GAAAGAATTTTTTTTCAGATATGTACTATCAAGGAACCTCGGATTAATACTAATACAA 186
DB 121 GAAAGAATTTTTTTTCAGATATGTACTATCAAGGAACCTCGGATTAATACTAATACAA 177
QY 187 GCTGATCATCAACACGTCATATCGGAATAGTAATAATAATAATAATTCAGACACTT 246
DB 178 GCTGATCATCAACACATCATATCATGGAATAGTAATAATAATAATAATTCAGACACTT 237
QY 247 TATTTGATGAACCTTAACATTTATATGCAAGGCTACACTACTTCTCACACA---CAGCAG 303
DB 238 TATTTGATGAACCTTAACATTTATATGCAAGGCTACACTACTTCTCACACAGCAGCAG 297

```

half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"

```
ORIGIN
Query Match      21.3%; Score 581.6; DB 2; Length 644;
Best Local Similarity 95.7%; Pred. No. 4.2e-126;
Matches 622; Conservative 0; Mismatches 19; Indels 9; Gaps 2;

QY 1229 GAGGAAATCGAAGGCTCAAGACTCAAAATTTGGGACCATCATCTAAGGCAACAAGCG 1288
    |||||
DB 1 GAGGAAATCGAAGGCTCAAGACTCAAAATTTGGGACCATCATCTAAGGCAACAAGCG 60
    |||||
QY 1289 CGTGCACAGATAGGAATGATGCAACCAATGCTTGGAGACCCCAAGAGGTTTACCTG 1348
    |||||
DB 61 CGCTGCAACAGTAGGAATGATGCAACCAATGCTTGGAGACCCCAAGAGGTTTACCTG 120
    |||||
QY 1349 AAGAGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1408
    |||||
DB 121 AAGAGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 180
    |||||
QY 1409 AGGATTCAGACAAATCATCTTGTCTAAGCAACCGGGCTAAACAAGGAGCCAGTGTCTA 1468
    |||||
DB 181 AAGACTCAGACAAATCATCTTGTCTAAGCAACCGGGCTTAAACAAGGAGCCAGTGTCTA 240
    |||||
QY 1469 ACTGGTTTCATAAATGCTTCGAGTTTCAATTTATGGAAGCCCAATGTTAGAGAAATGTTACTGG 1528
    |||||
DB 241 ACTGGTTTCATAAATGCTTCGAGTTTCAATTTATGGAAGCCCAATGTTAGAGAAATGTTACTGG 300
    |||||
QY 1529 AAGAGTGAAGATCAAGAACAAACAGTACTATCTTCTCAGGAGATACAAAAACAAG 1588
    |||||
DB 301 AAGAGTGAAGATCAAGAACAAACAGTACTATCTTCTCAGGAGATACAAAAACAAG 360
    |||||
QY 1589 AGACCAATATAAGTGTCTCCAAATGAAGAGAAACATCCAAATTTACTAGCAGCTTATTAC 1648
    |||||
DB 361 AGAGCAATATAGTGTCTCCAAATGAAGAGAAACACCAATTTATCTAGCAGCTTATTAC 420
    |||||
QY 1649 AAGATGGTATTACTACTCAAGCAGAAATTTCTACTCAACTATTTTCAACTTCCCTTA 1708
    |||||
DB 421 AAGAAGG-----TACTACTCAAGCAGAAATTTCTACTCAACTATTTTCAACTTCCCTTA 474
    |||||
QY 1709 CTGAGGTGCTTCACTTTCATCATGCTCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1768
    |||||
DB 475 CTGAGGTGCTTCACTTTCATCATGCTCAACACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 534
    |||||
QY 1769 ATATATCTACTACTACTGTTTGTATCATATTGAAACACGCGAAAGCAAGAA---ATG 1825
    |||||
DB 535 AGAATACTACTACTACTGTTTGTATCATATTGAAACACGCGAAAGCAAGCAAGCAATCATG 594
    |||||
QY 1826 ACATGCAAGTTTTCTCCAAAGTAGTATTTCTTTCATCTGTTGACATGGAA 1875
    |||||
DB 595 ACATGCAAGTTTTCTTCAAGTAATATTCTTTCATCTGTTGAAATGGAA 644
    |||||
```

```
RESULT 9
BG129304
LOCUS
DEFINITION EST474950 tomato shoot/meristem Lycopersicon esculentum cDNA clone
CTOP23N12 5' sequence, mRNA sequence.
ACCESSION BG129304
VERSION BG129304.1 GI:12629492
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 630)
van der Hoeven,R., Bezzerrides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished (2001)
CONTACT: CUGI
```

FEATURES

source

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.

Location/Qualifiers

1..630
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOP23N12"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/clone_lib="tomato shoot/meristem"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

ORIGIN

```
Query Match      20.7%; Score 565; DB 4; Length 630;
Best Local Similarity 95.4%; Pred. No. 3.4e-122;
Matches 606; Conservative 0; Mismatches 20; Indels 9; Gaps 2;

QY 1497 ATGGAAGCCCAATGGTAGAAGAAATGTACTTTGGAAGAGTGGAAGATCAAGAACAAACAG 1556
    |||||
DB 2 ATGGAAGCCCAATGGTTGAAGAAATGTACTTTGGAAGAGTGGAAGATCAAGAACAAACAG 61
    |||||
QY 1557 TACTAATACTTCAGGAGATACCAAAAACAAGACCAATATTAAGTGCTCCAATGAAGA 1616
    |||||
DB 62 TAGTAATACTTCAGGAGATACCAAAAACAAGACCAATATTAAGTGCTCCAATGAAGA 121
    |||||
QY 1617 GAAACATCAATTTATTACTAGCAGCTTATTACAAGATGGTATTACTACTACTCAAGCAGA 1676
    |||||
DB 122 GAAACCAACCAATTTATTACTAGCAGCTTATTACAAGATGG-----TACTACTCAAGCAGA 175
    |||||
QY 1677 AATTTCTACTCAACTATTTTCAACTTCCCTACTCGCAGGTGCTTCACTTCAATGCTCA 1736
    |||||
DB 176 AATTTCTACTCAACTATTTTCAACTTCCCTACTCGCAGGTGCTTCACTTCAATGCTCA 235
    |||||
QY 1737 CAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1796
    |||||
DB 236 CAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 295
    |||||
QY 1797 TGAACCAACGCGAAAAAGCAAGAA---ATGACATGCAACAGTTTCTCCAAGTAGTAT 1853
    |||||
DB 296 TGAACCAACGCGAAAAAGCAAGAAATCATGACATGCAACAGTTTCTCCAAGTAGTAT 355
    |||||
QY 1854 TCCTTTCATCTGTTGACATGGAAGCCAAAGCTAGAGAATCATCAATAAAGGGTTTACTAA 1913
    |||||
DB 356 TCCTTTCATCTGTTGAAATGGAAGCCAAAGCTAGAGAATCAACAAACAAAGGGTTTACTAA 415
    |||||
QY 1914 TCCTTTCATGCGCAGCATACGCGATGGGAGATTTTGGAGGTTTTCCTCTCATGATCAACA 1973
    |||||
DB 416 TCCTTTCATGCGCAGCATACGCGATGGGAGATTTTGGAGGTTTTCCTCTCATGATCAACA 475
    |||||
QY 1974 AATGACCGCAATTTTTCATGGAATTAATGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2033
    |||||
DB 476 AATGACCGCAATTTTTCATGGAATTAATGGTGTATCTCTCTCTCTCTCTCTCTCTCTCTCT 535
    |||||
QY 2034 TGAACCACTTAGCCATGCCAGTAGGCAACAAATAATACCTTTCTAATGACTTGGGAAGTAG 2093
    |||||
DB 536 TGAACCACTTAGCCATGCCAGTAGGCAACAAATAATATCTTTCTAATGAGTTGGGAAGTAG 595
    |||||
QY 2094 GTCTGAATGGGAGCATTCATTAATAGATGGAT 2128
    |||||
DB 596 GCCTGAATGAAGAGTCATTACAATAGAATGGAT 630
    |||||
```

RESULT 10
BI921008
LOCUS

BI921008 789 bp mRNA linear EST 10-MAR-2003

```
DEFINITION EST540943 potato microtubers, in vitro-grown Solanum tuberosum cDNA
clone cSTE27121 5' end, mRNA sequence.
ACCESSION BI921008
VERSION BI921008.1 GI:16217036
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 789)
van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R.,
Karamycheva,S.A., Tsai,J., Van Aken,S., Utterback,T., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generation of ESTs from in vitro grown microtubers (2001b)
Unpublished (2001)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.
FEATURES             Location/Qualifiers
     source
       1..789
         /organism="Solanum tuberosum"
         /mol_type="mRNA"
         /cultivar="Bintje"
         /db_xref="taxon:4113"
         /clone="cSTE27121"
         /tissue_type="axillary buds of stem explants; growing
         sink-tubers"
         /dev_stage="7, 8 and 10 days"
         /lab_host="SOLR"
         /clone_lib="potato microtubers, in vitro-grown"
         /notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
         XhoI. Tissue supplied by Christian Bachem and Richard
         Visser (Department of Plant Breeding, Wageningen
         University, The Netherlands); sequencing by the Institute
         for Genomic Research. The cDNA libraries will attempt to
         capture the induction and initiation/initial growth of the
         tuber in an in vitro system as described in Bachem et al.
         (Plant Journal, 1996). Small microtubers develop from
         axillary buds attached to stem explants when placed on a
         high sucrose medium (10%). Visible morphological changes
         occur synchronously at day five in the axillary buds. The
         first library, cSTA (1-20) consists of axillary buds
         harvested on days 1-3. This targets those genes involved
         in induction of the microtubers. The following libraries,
         cSTA (21-40) and cSTA (41-60), capture genes involved in
         tuber initiation and outgrowth. This library is noted as
         P3 in Tanksley lab notebooks."
ORIGIN
Query Match          20.6%; Score 564.2; DB 4; Length 789;
Best Local Similarity 99.5%; Pred. No. 5,6e-122;
Matches 566; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 87 TTACATGTTACTGTTATAGATAAAAGAGAAATCCGGAAGAAGAAATTTTTTTTCGAGA 146
|||
DB 221 TTGCAGGTACTGTTATAGATAAAAGAGAAATCCGGAAGAAGAAATTTTTTTTCGAGA 280
QY 147 TATGTTACTATCAGGAACCTCGGATAATACTAATATACAGCTGATCATCAACAGTCA 206
|||
DB 281 TATGTTACTATCAGGAACCTCGGATAATACTAATATACAGCTGATCATCAACAGTCA 340
QY 207 TAATCATGGAATAGTAATAATAATAATTTCAGACACTTTATTTGATGAACCTAACAA 266
|||
DB 341 TAATCATGGAATAGTAATAATAATAATAATTTCAGACACTTTATTTGATGAACCTAACAA 400
QY 267 TTATATGCAAGGTTACACTACTTCTGCACACACAGCAGCAGCAGTACTTTTCTCGAA 326
|||||
```

```
401 TTATATGCAAGGTTACACTACTTCTGCACACACAGCAGCAGCAGTACTTTTCTCGTGA 460
327 TTCTTACCAGCAGCAAGCAACCGCTTTGCCATGCGAATATACACACGCGCGCTGCA 386
|||||
461 TTCTTACCAGCAGCAAGCAACCGCTTTGCCATGCGAATATACACACGCGCGCTGCA 520
387 ACAGCAGCACTTTTGTGGGTGCTCTTCCGGCAGTAAGTTTGCAGCATCAGATCAATCA 446
521 ACAGCAGCACTTTTGTGGGTGCTCTTCCGGCAGTAAGTTTGCAGCATCAGATCAATCA 580
447 TCATGCACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACC 506
581 TCATGCACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACC 640
507 ATCGTCGACGGGGGTTTCTGCCACGTCTGTGCGCGGATCACCACGGACTTGGCGTCTCA 566
|||||
641 ATCGTCGACGGGGGTTTCTGCCACGTCTGTGCGCGGATCACCACGGACTTGGCGTCTCA 700
567 ATTGGCGTTTTCAGAGCGCGATTCCGACACCAACAACCGACAGCAACAACAGCAAGG 626
701 ATTGGCGTTTTCAGAGCGCGATTCCGACACCAACAACCGACAGCAACAACAGCAAGG 760
627 CGGTCTATCTTAAGCGCTTTTCTCTCAGC 655
|||||
761 CGGTCTATCTTAAGCGCTTTTCTCTCAGC 789

RESULT 11
CO909070
LOCUS BJ02039F01 BJ02 Capsicum annuum cDNA 5', mRNA linear EST 16-AUG-2004
DEFINITION CO909070
ACCESSION CO909070
VERSION CO909070.1 GI:51299373
KEYWORDS EST.
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Capsicum.
1 (bases 1 to 610)
Song,W.Y., Jeon,W.B., Kim,K.S., Lee,H.H., Ko,M.K., Kim,Y.S.,
Hong,J.C. and Oh,B.J.
Generation of Hot Pepper (Capsicum annuum) ESTs (Express Sequence
Tags) from Red Ripe Fruit (Song, et al.)
Unpublished (2004)
Contact: Boung-Jun Oh
Jeonnam Biotechnology Research Center
Namyang Bld. #603, 10-4 Gwangsan-dong, Dong-gu, Gwangju 501-180,
Korea
Tel: 82 62 607 2422
Fax: 82 62 607 6205
Email: bjo@biohub.re.kr
Plate: 039 row: F column: 01.
FEATURES             Location/Qualifiers
     source
       1..610
         /organism="Capsicum annuum"
         /mol_type="mRNA"
         /cultivar="Nokkwang"
         /db_xref="taxon:4072"
         /tissue_type="red pepper fruit pericarp"
         /clone_lib="BJ02"
         /note="Vector: pBluescript II SK(+); Site 1: EcoRI;
         Site 2: XhoI; cDNA library was generated from red ripe
         fruit pericarp using lambda Zap II phage vector. In vivo
         excision was done with helper phage to generate subclone
         in pBluescript II SK(+);XR vector."
ORIGIN
Query Match          19.4%; Score 531; DB 7; Length 610;
Best Local Similarity 92.5%; Pred. No. 3,6e-114;
Matches 558; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
QY 911 CTGGTGGTGGTGAAGTAGCAGCAGCAGCAAAAAATGAAGTTGCTGTTGAGCTTACAAC 970
```


AUTHORS

Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
Tanksey,S.D.

TITLE

Generation of ESTs from tomato fruit tissue, breaker stage

JOURNAL

Unpublished (2000)

COMMENT

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

source

Location/Qualifiers

1..721

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="cLEG25M5"

/tissue_type="Pericarp"

/dev_stage="breaker"

/lab_host="SOLR"

/clone_lib="tomato breaker fruit, TIGR"

/note="Vector: pBluescriptSKmCUadapt; Site 1: EcoRI;
Site 2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

ORIGIN

Query Match 18.4%; Score 502.4; DB 2; Length 721;
Best Local Similarity 88.6%; Pred. No. 2.1e-107;
Matches 654; Conservative 0; Mismatches 31; Indels 53; Gaps 8;

QY 2013 TACTTAGACATTCCTCTCTGAAACCTAGCCATGCCAGTGAGCCACAAAATTACCT 2072
DB 721 TACTTTAGGACTT-CTCTTTCTGAAACCTAGCCATGCCAGTGAGCCACAAA-TATCT 664
QY 2073 TTTCTAATGACTGGGAAGTAGGCTCAAAATGGGAGTCATTACAATAGATGGGATATCA 2132
DB 663 TTTCTAATGAGTTGGGAAGTAGGCTCAAAATAGGAAGTCTATTACAATAGATGGGATATGA 604
QY 2133 AAACATTGATTTTCAGAGTGGGAATAGCGATTTCCGACTCACTATTACCGAGATTTTGT 2192
DB 603 AAACATTGATTTTCAGAGTGGGAATAGCGATTTTCGAACTCACTATTGCCAGATTTTGT 544
QY 2193 TACAGGTATCTAGGAACATGAATACAGAAAGTCTCGTATTGATAGCTGAAAAGAT--A 2250
DB 543 TACAGGTATCTTGGAAACATGAATACCAAAAGTCTCGTATTGATAGCTGAAAACATAAA 484
QY 2251 AAAGGAAGTTAGGGTACTCTTATATTGTGTAGGCTTCTGSCCCCAAGTCGGAGGACCC 2310
DB 483 AAAGGAAGTTAGGGTACTCTTATATTGTGTAGGCTTCTGSCCCCAAGTCGGAGGACCC 424
QY 2311 AATTTTGATACAACTATCATAGGAAAAAGAGTGGAGACTAAATTTAAAGTAAACAAAT 2370
DB 423 AATTTTGATACAACTATCATAGGAG-----ACAAAGT 391
QY 2371 TTAACGACACTTCTAGTATATATATATCTCTTTT---TTATAGTATAGAAAAGAGAG 2427
DB 390 TTAAGCAGACTTCTAGTATATATATATCTCTTTT---TTAATATAGTATAGAAAAGAGAG 331
QY 2428 ATTTTGTCTTTAGTCTATAGATAGAGTCTACTTAGTAGGTTACTTCTAGTTCCTT 2487
DB 330 ATTTTGTCTTTAGTCTATAGATAGAGTCTACTTAGTATAGGTTACTTCTACTTCTT 271
QY 2488 GAGAAGATTGATCAAACTAGTAGTATTTTTCCTTTTGG-----GTTCGGCTTGG 2537
DB 270 GAGAAGATTGATCAAACTAGTAGTATTTTATTTTGGGGGGTTGAATAGCTTGTCTTGG 211
QY 2538 AGTACTATTTTAGTTAT-----TGGAACTAGCTATAG--TAATGTTGTAAGTTG 2588
DB 210 AGTACTATTTTAGTTAATGGAATTTTAAACTAGCTATAGTATATATGTTGTAAGTTG 151

QY 2589 TGATATTGTTCTCTCAATTTGCATATATATTTGAAATATTTTGTACCTACTAGTAGTCT 2648
DB 150 TGATATTGTTCTCTCAATTTGCATATATATTTGAAATATTTTGTACCTACTAGTAGTCT 91
QY 2649 CTAATATTGTTTCCATTGCTTGAATGCAATTTTATTTGAATTTTGTGCTATCATAT 2708
DB 90 CTAATATTGTTTCCATTGCTTGAATGCAATTTTATTTGAATTTTGTGCTATCATAT 31
QY 2709 TAGATTAGCAAAAAA 2726
DB 30 TAGATTAGCAAAAGTACA 13

RESULT 14
CK262102
LOCUS
DEFINITION
CK262102 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POABE14 5' end, mRNA sequence.
CK262102
VERSION
CK262102.1 GI:39819080
KEYWORDS
SOURCE
Solanum tuberosum (potato)
ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 964)
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: EST708181
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>
Seq primer: ANT TAG GTG ACA CTA TAG.

FEATURES

source

1..964

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="POABE14"

/tissue_type="abiotic stress treated leaf and root tissue"

/lab_host="DH10B-TonA"

/clone_lib="potato abiotic stress cDNA library"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: Solanum tuberosum var. Kennebec plants were

grown from cuttings on a 16hr light/8 hr dark cycle at 25

C for 3-4 weeks. Abiotic stress conditions were applied to

four separate sets of plants. Set 1 involved saturation of

the soil with 150 mM NaCl and tissues were harvested at

following application of the salt stress (leaves: 2hr,

6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).

Set 2 were grown under the standard conditions and then

were water stressed by withdrawal of further watering

applications. Drought stressed plants were harvested after

cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d

and 5d). Set 3 were grown under the standard conditions

and then were cold stressed by placement at 4 C. Cold

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,

2d. Set 4 were grown under the standard conditions and

then were heat stressed by placement at 35 C. Heat

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

2d and 4d and heat-stressed roots were harvested at 6 hr,

12 hr, 1 d, and 4d. RNA was isolated from all tissues and

equal RNA from each tissue and stress was pooled to

construct the cDNA library. RNA sample."

ORIGIN

Query Match	17.4%;	Score 476;	DB 7;	Length 964;	
Best Local Similarity	72.6%;	Pred. No. 3.8e-101;			
Matches 687;	Conservative	0;	Mismatches 220;	Indels 39;	Gaps 4;
QY	909	TTCTGGTGGTGAAGTAGCAGCAGGAGAAATAATGAAGTTGCTGTGAGCTTCAAC	968		
DB	34	TTTGGCTAGTGTGATGTTGAAAGTAGCAGCCAAATAATTTGTTGTAACCTTACTAC	93		
QY	969	TGCTCAAAGACAAGACTTCAATGAAAGGCAAGCTTCTTGCCATGCTTGAAGAGGT	1028		
DB	94	AGTCAAAGACAAGACTTCAATGAAAGGCAAGCTTGTAGCATGCTTGAAGAGGT	153		
QY	1029	GGAGCAAAGGTACAGACAGTACCATCAACCAATGCAAAATATGTAATATCATTTGAGCA	1088		
DB	154	GGATCAAAGGTATAGACAATACCATCACCAATGCAAAATGATGCAACATCATTTGAGCA	213		
QY	1089	AGTAGAGGAATGGATGACGCCAAATCATACACTCAATAGCTTTGATGCAATTTGAA	1148		
DB	214	AACAACAGGAATGGATCATCAAAATCATACACAACCTTGTGTCACACAATTTGAA	273		
QY	1149	GCAATTCAGATGCTAAAGAGTCAATTTGCTGAGCAAGTAAAGGCGACGACGAGAGTTT	1208		
DB	274	GCAATTTAGATGTTTAAAGATGCAATTTCTGGCAATTAAGACACAAGTAAACCTTT	333		
QY	1209	AGGTGAAGAGGAGGCTTGGGAGGGAATAATCGAAGGCTCAAGACTCAAAATTTGTGACCA	1268		
DB	334	AGGGGAAGAGAGACAATTTGGAGGCAAAATTTGAAGGATCAAGTTGAAATTTGTGATCA	393		
QY	1269	TCATCTAAGCAACAACGCGGCTGCAACAGATAGGAATGATGAACCAATGCTTGGAG	1328		
DB	394	TCATTTACGCAACAACGCTGCACTACACAATTAGGGATGATGCAACCAATCATGGAG	453		
QY	1329	ACCCAAAGAGGTTTACTGAAGAGTGTCTGTCTCTGCTGCTGCTGCTTTTCGAGCA	1388		
DB	454	GCAACAGAGGTTTGGCCGAAGAGGCGTTTGGTTCTCCGCGCTTGGCTTTTCGAGCA	513		
QY	1389	TTTTCTTCATCTTACCCAAGGATTCAGACAATAATCATGCTTGTAAAGCAAAACGGGCT	1448		
DB	514	TTTTCTTCATCTTACCCAAGGATTCAGACAATAATCATGCTTGTAAAGCAAAACGGGCT	573		
QY	1449	AACAAGAGCAGGTGTCTAATGTTTCATTAATGCTCGAGTTGATATGGAAGCCAAT	1508		
DB	574	AACAAGAGCAGGTATCAAAATGGTTTATAAATGCTAGAGTTAGACTATGGAAGCCAAT	633		
QY	1509	GGTAGAAGAAATGTACTTGAAGAGTGAAGATCAAGAACAAACAGTACTAATCTTC	1568		
DB	634	GGTAGAAGAAATGTACATGGAAGAGTGAAGAAACCAATCAAGACAATAATTTAGCC	693		
QY	1569	AGGAGATAACAACAAAGAGACCAATATAAGTGTCTCCAAATCAAGAGAAACATCCAAT	1628		
DB	694	TAATAACAATGAAATTTGTTGTTCAAAATCAAGTTTCCA--CAAGAGAAATTACCAAT	750		
QY	1629	TATTACTAGCAGCTTATTAACAAGTGTATTACTACTACTCAAGCAGAAATTTCTACCTC	1688		
DB	751	TAGTAGCA-----ATATTATTATTAATGCTTCTCCAAATGATATTTCTACTTC	798		
QY	1689	AACATTTTCACTTCCCTACTGCGGTGCTTCACTTCATCATGCTCTCACA-----A	1739		
DB	799	CACATTTTCAACATCTCCGACAGCGCGGGGTTCGATTCGGCTCAGACGGTTGAGG	858		
QY	1740	TTTCTCTCTCTTGGTTTCACTCAACATGGATAATTACTACTACTTGTGATCATATGA	1799		
DB	859	TTTCTCTCTTCACTTAGGTTCATTAACATGGAGACA-----TTGATGATCA	903		
QY	1800	AAACAACGCGAAAGCAAGAAATGACATGCAAGTTTCTCCA	1845		
DB	904	AAGGAACAACAAGGCAAGAAATGAGATGCAAAATTTGTTCACTA	949		
RESULT 15					
BE460264					
LOCUS					
		495 bp	mRNA	linear	EST 18-MAY-2001

DEFINITION	EST411755 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone cLEG28B4, mRNA sequence.				
ACCESSION	BE460264				
VERSION	BE460264.1 GI:9504638				
KEYWORDS	EST.				
SOURCE	Lycopersicon esculentum (tomato)				
ORGANISM	Lycopersicon esculentum				
REFERENCE	1 (bases 1 to 495)				
AUTHORS	Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tankalev,S.D.				
TITLE	Generation of ESTs from tomato fruit tissue, breaker stage				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html 5 prime sequence.				
FEATURES	Location/Qualifiers				
source	1..495				
	/organism="Lycopersicon esculentum"				
	/mol_type="mRNA"				
	/cultivar="TA496"				
	/db_xref="taxon:4081"				
	/clone="cLEG28B4"				
	/tissue_type="Pericarp"				
	/dev_stage="breaker"				
	/lab_host="SOLR"				
	/clone_lib="tomato breaker fruit, TIGR"				
	/notes="Vector: pBluescriptSKmCudapt; Site 1: EcoRI; Site_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom and the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."				
ORIGIN					
Query Match	17.1%;	Score 467.8;	DB 2;	Length 495;	
Best Local Similarity	96.6%;	Pred. No. 2.9e-99;			
Matches 478;	Conservative	0;	Mismatches 17;	Indels 0;	Gaps 0;
QY	1042	AGACAGTACCATCAACCAATGCAAAATTAATGTTATTCATTTCAGCAAGTAGCAGGAATT	1101		
DB	1	AGACAGTCCCATCCCAATGCAAAATTAATGTTATTCATTTCAGCAAGTAGCAGGAATT	60		
QY	1102	GGATCAGCCAAATCATACACTCAATTTAGCTTTGCATGCAATTTTGAAGCAATTCAGATGC	1161		
DB	61	GGATCAGCAAAATCATACACTCAATTTAGCTTTGCATGCAATTTTGAAGCAATTCAGATGC	120		
QY	1162	CTAAAGATGCAATTTGCTGAGCAAGTAAGGCGACGAGCAAGAGTTTAGTGAAGAGAA	1221		
DB	121	CTAAAGATGCAATTTAGTGAAGTAAGGCGACGAGCAAGAGTTTAGTGAAGAGAA	180		
QY	1222	GGCTTGGAGGGGAAATCGAAGGCTCAAGACTCAAAATTTGTGACCATCATCTTAAGGCAA	1281		
DB	181	GGCTTGGAGGGGAAATCGAAGGCTCAAGACTCAAAATTTGTGACCATCATCTTAAGGCAA	240		
QY	1282	CAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAAATGCTTTGGAGACCCCAAGAGGT	1341		
DB	241	CAACGCGCGCTGCAACAGCTAGGAATGATGCAACCAAAATGCTTTGGAGACCCCAAGAGGT	300		
QY	1342	TTACCTGAAAGAGCTGTCTGCTTCCTTCGCTTCGCTTCGCTTCGAGCAATTTCTTCATCCT	1401		
DB	301	TTACCTGAAAGAGCTGTCTGCTTCCTTCGCTTCGCTTCGCTTCGAGCAATTTCTTCATCCT	360		
QY	1402	TACCCAAAGGATTTCAGACAAATCATGCTTGTGCTTAAGCAACGCGGCTTAACAGAGCCAG	1461		
DB	361	TACCCAAAGGATTTCAGACAAATCATGCTTGTGCTTAAGCAACGCGGCTTAACAGAGCCAG	420		

Qy	1462	GTCTCTAACTGGTTCTATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATG	1521
Db	421	GTCTCTAACTGGTTCTATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATG	480
Qy	1522	TACTTGGAAAGAGTG	1536
Db	481	TACTTGGAAAGAGTG	495

Search completed: July 29, 2005, 14:13:05
Job time : 8595 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 11:50:07 ; Search time 8001 Seconds
(without alignments)
4166.628 Million cell updates/sec

Title: US-10-624-201a-2

Perfect score: 3563

Sequence: 1 MYQGSTDNITQADHQRRH.....GNKRFTQLLPDFVTGNLGT 688

Scoring table:

	BLOSUM62	Xgapext	0.5
Xgapop 10.0 , Xgapext	0.5		
Xgapop 10.0 , Xgapext	0.5		
Fgapop 6.0 , Fgapext	7.0		
Delop 6.0 , Delext	7.0		

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10624201/runat.27072005.122957.10905/app.query.fasta_1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US10624201 @CGN 1 1 5081 @runat.27072005.122957.10905 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3563	100.0	2735	8 AF406697	AF406697 Solanum t
2	3294.5	92.5	2636	8 AF375966	AF375966 Lycopersi
3	3294.5	92.5	2687	8 BT013459	BT013459 Lycopersi
4	1332	37.4	1898	8 AF406698	AF406698 Solanum t

5	1150.5	32.3	2043	6	AX506642	Sequence
6	1150.5	32.3	2043	8	BT008422	Arabidops
7	1150.5	32.3	2074	8	AY096548	Arabidops
8	1150.5	32.3	2572	8	AY072175	Arabidops
9	1150.5	32.3	2578	8	AY098854	Arabidops
10	1148.5	32.2	2385	8	AF353094	Arabidops
11	1127.5	31.6	2128	8	AF406702	Solanum t
12	999	28.0	120938	8	AC007017	Arabidops
13	967.5	27.2	2489	8	AK120844	Oryza sat
14	946	26.6	2643	8	AK120813	Oryza sat
15	945	26.5	3030	8	AK070465	Oryza sat
16	931.5	26.1	2065	8	AF406703	Solanum t
17	928	26.0	1896	8	AK104872	Oryza sat
18	926.5	26.0	2630	8	AK069994	Oryza sat
19	922.5	25.9	2523	8	AY524972	Oryza sat
20	911.5	25.6	2570	8	AK067939	Oryza sat
21	898.5	25.2	2476	8	AK121356	Oryza sat
22	891.5	25.0	2952	8	AK070543	Oryza sat
23	876	24.6	28712	8	AC145458	Cicer ari
24	839.5	23.6	2636	8	AK068034	Oryza sat
25	837.5	23.5	2698	8	AK099829	Oryza sat
26	835.5	23.4	2414	8	AK105752	Oryza sat
27	826	23.2	2050	8	AF353092	Arabidops
28	825.5	23.2	1906	8	AY570508	Arabidops
29	824	23.1	1884	6	AX506210	Sequence
30	824	23.1	1884	8	AY142019	Arabidops
31	824	23.1	2124	8	AY058179	Arabidops
32	819	23.0	1946	8	AY598452	Arabidops
33	819	23.0	1973	8	AY062545	Arabidops
34	816.5	22.9	146585	8	CNS08CAM	Arabidops
35	809	22.7	1950	8	AF353093	Arabidops
36	803	22.5	1824	8	AY085278	Arabidops
37	802	22.5	1449	6	AX412914	Sequence
38	802	22.5	1599	6	AX507459	Sequence
39	802	22.5	1599	8	BT012291	Arabidops
40	802	22.5	1634	8	BT000133	Arabidops
41	797.5	22.4	2545	8	AF173816	Arabidops
42	797.5	22.4	3162	8	AY050459	Arabidops
43	797	22.4	1449	8	BT000609	Arabidops
44	797	22.4	2126	8	AY056796	Arabidops
45	793.5	22.3	3132	8	GGN318871	Gnetum gn

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED
AF406697	Solanum tuberosum	BE1-related homeotic protein 5 (Bel5) mRNA, complete cds.	AF406697	AF406697.1	GI:22652114	Solanum tuberosum (potato)	Solanum tuberosum	Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.	Chen,H., Rosin,F.M. and Hannapel,D.J.	A KNX protein of potato interacts with several members of the TALE extension superclass regulate tuber formation	Plant Physiol. 132 (3), 1391-1404 (2003)	12857821	2 (bases 1 to 2735)	Chen,H., Rosin,F.M. and Hannapel,D.J.	A KNX protein of potato interacts with several members of the TALE extension superclass regulate tuber formation	Plant Physiol. 132 (3), 1391-1404 (2003)	12857821	3 (bases 1 to 2735)	Chen,H., Rosin,F.M. and Hannapel,D.J.	Direct Submission	Unpublished	Unpublished
AF406697	Solanum tuberosum	BE1-related homeotic protein 5 (Bel5) mRNA, complete cds.	AF406697	AF406697.1	GI:22652114	Solanum tuberosum (potato)	Solanum tuberosum	Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.	Chen,H., Rosin,F.M. and Hannapel,D.J.	A KNX protein of potato interacts with several members of the TALE extension superclass regulate tuber formation	Plant Physiol. 132 (3), 1391-1404 (2003)	12857821	2 (bases 1 to 2735)	Chen,H., Rosin,F.M. and Hannapel,D.J.	A KNX protein of potato interacts with several members of the TALE extension superclass regulate tuber formation	Plant Physiol. 132 (3), 1391-1404 (2003)	12857821	3 (bases 1 to 2735)	Chen,H., Rosin,F.M. and Hannapel,D.J.	Direct Submission	Unpublished	Unpublished
AF406697	Solanum tuberosum	BE1-related homeotic protein 5 (Bel5) mRNA, complete cds.	AF406697	AF406697.1	GI:22652114	Solanum tuberosum (potato)	Solanum tuberosum	Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.	Chen,H., Rosin,F.M. and Hannapel,D.J.	A KNX protein of potato interacts with several members of the TALE extension superclass regulate tuber formation	Plant Physiol. 132 (3), 1391-1404 (2003)	12857821	2 (bases 1 to 2735)	Chen,H., Rosin,F.M. and Hannapel,D.J.	A KNX protein of potato interacts with several members of the TALE extension superclass regulate tuber formation	Plant Physiol. 132 (3), 1391-1404 (2003)	12857821	3 (bases 1 to 2735)	Chen,H., Rosin,F.M. and Hannapel,D.J.	Direct Submission	Unpublished	Unpublished

University, 257 Horticulture Hall, Ames, IA 50011-1100, USA

FEATURES

source

1..2735

Location/Qualifiers

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Desiree"

/db_xref="taxon:4113"

1..2735

/gene="Bel5"

148..2214

/gene="Bel5"

/note="StBel5"

/codon_start=1

/product="BEL1-related homeotic protein 5"

/protein_id="AA03621.1"

/db_xref="GI:22652115"

/translation="MYQGTSDNTNTIQADHQHNGNSNNNIOTLYLMNPNYMQG
YTTSDTQQOQLFLNSSPAASNALCHANOHPAQHQHFGVGPVPAVSLHLDQINHHG
LLQRMNNQDQSQVTPSGTGVTSATSCGGITTDLSAQAFORPPTPQHRQOQOQOQ
GLSLSPLOQOQISFNNNISSSPNTNVTIRGLDSSSNMVLGSKYLKAAQELLQ
EVNIVKSIKQDDOKDNMKNESPLASDVNTSSGGESSRQKNIEVAVELTIAQ
RQELQKKAKLMLKEVQRIRQYHQHQIIVLSFEQVAGIGSAKSYTQIALHAISK
QFRLCKDLAEQKATSKSLGEEGLGKIEGRSLKVDHLRQORALQOIGMMQFNA
WRPQRLPERAVSVLRAWLFEHLPHYPKDSKIMLAKOTGLTRSOVSNWFNARVL
WPKPMVEMLYEEVKNQOENSTNTSGDNKNETNISAPNEEKPIITSSLLQDGIITVQ
ABISTSTPTAGASLHHAHNFPLGSFNMNDTTTVDHIENNAKKORNDMHKFSF
SSILSSVDMBAKAESNKGFTNPLMAAYAMGDGRFPDHDQOQMTANFHNNGVSLTL
GLPPSENLPVSSQQNLYLNDLGSRSSEMSHYNRMGYENIDFQSGNKRFPQTQLLPDFV
TGNLGT"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

1,116-249

Length:

2735

3563.00

Matches:

688

100.00%

Conservative:

0

100.00%

Mismatches:

0

100.00%

Indels:

0

8

Gaps:

0

US-10-624-201A-2 (1-688) x AF406697 (1-2735)

QY

1

MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnArgHis

20

DB

148

ATGTACTATCAAGAACTCGGATAATACTAATATACAAGCTGATCATCAACACGTCAT

207

QY

21

AsnHisGlyAsnSerAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn

40

DB

208

AATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTGTGATGAACCCCTAACAA

267

QY

41

TyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGlnLeuLeuPheLeuAsn

60

DB

268

TATATCGAAGGCTACTACTTCTGACACACAGCAGCAGCAGTACTTTTCTCGAAT

327

QY

61

SerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGln

80

DB

328

TCCTTACCAGCAGCAAGCAACGGCTTTGCGCATGCGAATATACAACACGCGCGCTGCAA

387

QY

81

GlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsnHis

100

DB

388

CAGCAGCACTTTGCGGTGCGCTCTTCGGCGCAGTAAGTTTGACCATCAGATCATCATCAT

447

QY

101

HisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValPro

120

DB

448

CATGGACTTTTACAGCGCATGTGGACACCAACCAAGATCAATCTCAGCAGGTGATAGTACCA

507

QY

121

SerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGln

140

DB

508

TCGTCACGGGGGTTTCTGCCAGCTCATGTGGCGGATCACCACGACTTGGCGTCTCAA

567

QY

141

LeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGly

160

DB

568

TTGGCGTTTTCAGAGGCGGATTCGACACCAACACCGACGAGCAACACGCAAGGC

627

QY	161	GlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIle 180
DB	628	GGTCTATCTTAAGGCTTCTCTCTCAGCTACACAGCAAAATTAGTTTCAATAACAATATT 687
QY	181	SerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySerSer 200
DB	688	TCATCTCTCATCAAGCAAGCAAAATAATGTACTATTAGGGGAACATTAGATGAAGTTCT 747
QY	201	SerAsnMetValLeuGlySerIlySerIlyValLeuYsAlaAlaGlnGlnLeuLeuAspGluVal 220
DB	748	AGCAACATGGTTTTTAGGCTCTAAGTATCTGAAGGCTGCAAGAGAGCTTCTTATGAAGTT 807
QY	221	ValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsn 240
DB	808	GTTAATATTGTTGGAAAGAGCATCAAGAGGATGATCAAAAGAGGATTAATTCATGAAT 867
QY	241	LysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSer 260
DB	868	AAAGAATCAATGCCCTTTGGCTAGTGTCAACACTAATAGTTCTGGTGGTGAAGT 927
QY	261	SerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGlnLeu 280
DB	928	AGCAGCAGCAGCAAAATGAAGTTGCTGTGAGCTTACAACCTGCTCAAGAGCAAGAACTT 987
QY	281	GlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGln 300
DB	988	CAATGAAAGAAAGCCAACTTCTGCCATGCTTTGAAGAGGTGAGCAAGGTTACAGACAG 1047
QY	301	TyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySer 320
DB	1048	TACCATCAACAAATGCAAAATAATTGTTATTATCATTTGAGCAAGTAGCAGGAATTTGGATCA 1107
QY	321	AlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLys 340
DB	1108	GCCAAATCATACACCAATTAGCTTTGCAATGCAATTTTGAAGCAATTCAGATGCCCTAAAG 1167
QY	341	AspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGlyLeu 360
DB	1168	GATGCAATTTGCTGAGCAAGTAAAGCGCAGCAGCAAGAGTTTAGGTCAAGAGGAGGCTTG 1227
QY	361	GlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnArg 380
DB	1228	GGAGGGAAATCGAAGGCTCAAGACTCAAAATTTGTGGACCATCATCTAAGGCAACAAACG 1287
QY	381	AlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuPro 400
DB	1288	GGCTGCAACACATAGGAATGATGCAACCAATGCTTGGAGACCCCAAGAGGTTTACCT 1347
QY	401	GluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrPro 420
DB	1348	GAAAGAGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1407
QY	421	LysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSer 440
DB	1408	AAGGATTCAGACAAATCATGCTTGTCTAAGCAAAACGGGGCTTAACAGGAGGAGGTTGCT 1467
QY	441	AsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeu 460
DB	1468	AACTGTTTCAATAATGCTCGAGTTGATTTATGGAAGCCAAATGTTAGGAAGAAATGACTTG 1527
QY	461	GluGluValLysAsnGlnGlnGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
DB	1528	GAAGAAGTGAAGATCAAGAACAAACACTACTAATACTTTCAGGAGATTAACAAAACAAA 1587
QY	481	GluThrAsnIleSerAlaProAsnGlnGluLysHisProIleIleThrSerSerLeuLeu 500
DB	1588	GAGACCAATATAGTGTCTCAATGAAGAGAAACATCCAAATTTATTACTAGCAGCTTATTA 1647
QY	501	GlnAspGlyIleThrThrGlnAlaGluIleSerThrSerThrThrIleSerThrPro 520
DB	1648	CAAGATGTTATTACTACTACTACTCAAGCAGAAATTTCTACCTCAACTATTTCACCTCCCT 1707
QY	521	ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540

QY	268	ValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuMetIysLysAlaLysLeu	287
Db	937	:::	996
QY	288	LeuAlaMetLeuGluValGluGlnArgTyrArgGlnTyrHisGlnMetGlnIle	307
Db	997		1056
QY	308	IleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeu	327
Db	1057	:::	1116
QY	328	AlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnVal	347
Db	1117	:::	1176
QY	348	LysAlaThrSerLysSerLeuGlyGluGluGlyLeuGlyLysIleGluGlySer	367
Db	1177	:::	1236
QY	368	ArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnIleGlyMet	387
Db	1237		1296
QY	388	MetGlnProAsnAlaIleTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeu	407
Db	1297		1356
QY	408	ArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMet	427
Db	1357		1416
QY	428	LeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArg	447
Db	1417		1476
QY	448	ValArgLeuTrpLysProMetValGluGluMetTyrLeuGluValLysAsnGlnGlu	467
Db	1477		1536
QY	468	GlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaPro	487
Db	1537		1596
QY	488	AsnGluGluLysHisProIleThrSerSerLeuLeuGlnAspGlyIleThrThrThr	507
Db	1597		1650
QY	508	GlnAlaGluIleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis	527
Db	1651		1710
QY	528	HisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrVal	547
Db	1711		1770
QY	548	AspHisIleGluAsnAsnAlaLysLysGlnArgAsn--AspMethIlySypheSerPro	566
Db	1771		1830
QY	567	SerSerIleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGly	586
Db	1831		1890
QY	587	PheThrAsnProLeuMetAlaIleTyrAlaMetGlyAspPheGlyArgPheAspProHis	606
Db	1891		1950
QY	607	AspGlnGluMetThrIleAsnPheHisGlyAsnAsnGlyValSerLeuThrIleGlyLeu	626
Db	1951		680
QY	627	ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu	646
Db	2011		7070
QY	647	GlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGluAsnIleAspPhe	666
Db	2071		7130
QY	667	GlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheValThrGlyAsnLeu	686
Db	2131		7190
QY	687	GlyThr 688	
Db	2191	GGAAACA 2196	
RESULT 3			
BT013459		2687 bp mRNA linear PLN 11-MAY-2004	
LOCUS		Lycopersicon esculentum clone 132117F, mRNA sequence.	
DEFINITION		BT013459	
ACCESSION		BT013459	
VERSION		BT013459.1 GI:47104874	
KEYWORDS		FLI_CDNA.	
SOURCE		Lycopersicon esculentum (tomato)	
ORGANISM		Lycopersicon esculentum	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;	
TITLE		asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.	
JOURNAL		1 (bases 1 to 2687)	
FEATURES		Kirkness,E.F., Wang,W. and Vazeille,A.	
source		Submitted (11-MAY-2004) The Institute for Genomic Research, 9712	
		Medical Center Drive, Rockville, MD 20850, USA	
		Location/Qualifiers	
		1..2687	
		/organism="Lycopersicon esculentum"	
		/mol_type="mRNA"	
		/db_xref="taxon:4081"	
		/clone="132117F"	
		/tissue_type="fruit"	
		/note="TZXBC46"	
ALIGNMENT SCORES:			
Pred. No.:	3 8e-230	Length:	2687
Score:	3294.50	Matches:	647
Percent Similarity:	95.16%	Conservative:	21
Best Local Similarity:	92.17%	Mismatches:	17
Query Match:	92.46%	Indels:	7
DB:	8	Gaps:	7
US-10-624-201A-2 (1-688) x BT013459 (1-2687)			
QY	1	MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAsp---HisGlnGlnArg	19
Db	154	ATGTACTACTCAAGGAACCTCGGATAAT---AATATACAAGCTCATCATCAACAACAA	210
QY	20	HisAsnHisGlyAsnSerAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsn	39
Db	211	CATAATAATCTTGGGAATAGTAATAATAATATTCAGACGCTGTATCTGATGAACCC	270
QY	40	AsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGln-----	52
Db	271	AGTTATATGCAAGGCTACACTACACTGCACACAGCAGCAGCTGCACAGCAACAGAAC	330
QY	53	GlnGlnGlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAla	72
Db	331	CAGCACCAGTCTCTTTCTGAATTCAGCACCAGGAGGCAACGCGCTTAGCCATGCG	390
QY	73	AsnIleGlnHisAlaProLeuGlnGlnHisPheValGlyValProLeuProAlaVal	92
Db	391	AATATACAACACGCGCTCTGCAACAGCAGCAGCTTTGTGCGCGTCCCTTCGCGCAGTG	450

Qy 93 SerLeuHisAspGlnIleAenHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAsp 112
 |||||
 Db 451 AGTTTGCATGATCAGATCAATCATCAGGACTTTTACAGCGTATGTGGAACAATCAAGAC 510
 |||||
 Qy 113 GlnSerGlnGlnValIleValProSerSerThrClyValSerAlaThrSerCysGlyGly 132
 |||||
 Db 511 CAATCTCAGCAGGAGTAGTACCGTCTCGACGGTGTCTTCTGCCACGTATGTGGCGGG 570
 |||||
 Qy 133 IleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIle-----ProThr 149
 |||||
 Db 571 ACCACCAAGGACTTGGCGTCTCAACTGGCGTTTCAGAGGCGGATTTGGTGTCTCCGACA 630
 |||||
 Qy 150 ProGlnHisArgGlnGlnGlnGlnGlnGlnGlyLeuSerLeuSerLeuSerProGln 169
 |||||
 Db 631 CCACAACCCGACACAACAACAACAACAAGGCGTCTATCACTGAGCCTTTCTCTCTCAG 690
 |||||
 Qy 170 LeuGlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerProArgThrAsnAsn 189
 |||||
 Db 691 CAACAGCACAATATTAGTTTCAATTAACAATATTTCATCATCATCACCAGGACAATAAT 750
 |||||
 Qy 190 ValThrIleArgGlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLysTyr 209
 |||||
 Db 751 GTTACTATTAGAGGAACAATGGATGTTCTAGCAACATGATTTTAGGCTCTAAGTAT 810
 |||||
 Qy 210 LeuLysAlaAlaGlnGlnLeuLeuAspGluValValAsnIleValGlyLysSerIleLys 229
 |||||
 Db 811 CTGAAAGCTGCACAAGAGCTTCTTGATGAAGTGGTTAATATTGTTGAAAAAGTAACAAA 870
 |||||
 Qy 230 GlyAspAspGlnLysLysAspAsnSerMetAsnLysGluSerMetProLeuAlaSerAsp 249
 |||||
 Db 871 GGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAAATTAATTCCTTTGGTTAGTGAT 930
 |||||
 Qy 250 ValAsnThrAsnSerSer-----GlyGlyGlyGluSerSerSerArgGlnLysAsnGlu 267
 |||||
 Db 931 GTCAATACTAATAGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 990
 |||||
 Qy 268 ValAlaValGluLeuThrAlaGlnArgGlnGlnLeuGlnMetLysLysAlaLysLeu 287
 |||||
 Db 991 GTTGCCATTGAGCTTCAACTGCTCAAAAGACAAGAACTTCAAAATGAAAAAGCAAGCTT 1050
 |||||
 Qy 288 LeuAlaMetLeuGluValGluGlnArgTyrArgGlnTyrHisGlnMetGlnIle 307
 |||||
 Db 1051 CTTGCCATGCTTGAAGAGGTGGAGCAAGGTACAGACAGTACCATCAACCAATGCAATA 1110
 |||||
 Qy 308 IleValLeuSerPheGluGlnValAlaGlyLeuGlySerAlaLysSerTyrThrGlnLeu 327
 |||||
 Db 1111 ATGTATCATCAATTGAGCAAGTAGCAGAGGATGGATCAGCAAAATCATACACTCAATTA 1170
 |||||
 Qy 328 AlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnVal 347
 |||||
 Db 1171 GCTTTGCATGCAATTTCAAGCAATTCAGATGCCCTAAGGATGCAATTTAGTGACCAAGTG 1230
 |||||
 Qy 348 LysAlaThrSerLysSerLeuGlyGluGluGlyLeuGlyGlyLysIleGluGlySer 367
 |||||
 Db 1231 AAGCGCAGCAGCAGAGAGTTAGGTGAAGATGAAGCTTGGAGGAAAAATTTGAAGGCTCA 1290
 |||||
 Qy 368 ArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMet 387
 |||||
 Db 1291 AGACTCAAAATTTGTGGACCATCATCTAAGGCACAACCGCGCTGCAACAGCTAGGAATG 1350
 |||||
 Qy 388 MetGlnProAsnAlaIleTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeu 407
 |||||
 Db 1351 ATGCACCAAAATGCTTGAGACCCCAAGAGGTTTACCTGAAAGAGCTGTCTGTCTCTT 1410
 |||||
 Qy 408 ArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMet 427
 |||||
 Db 1411 CGTGTCTGGCTTTTTCGAGCATTTTCTTCATCTTACCTTACCCCAAGACTCAGACAAATCATG 1470
 |||||
 Qy 428 LeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArg 447
 |||||
 Db 1471 CTTGCTAAGCAACAGGGGCTTAACAGGAGCCAGGCTCTTAAGTGTTCATTAATGCTCGA 1530
 |||||
 Qy 448 ValArgLeuTrpLysProMetValGluGluMetTyrLeuGluValLysAsnGlnGlu 467
 |||||

Db 1531 GTTCGATTATGGAAGCAATGGTTGAAGAAATGTACTTGAAGAAGTGAAGAATCAAGAA 1590
 |||||
 Qy 468 GlnAsnSerThrAsnThrSerGlyAspAsnLysGlnThrAsnIleSerAlaPro 487
 |||||
 Db 1591 CAAAACAGTAGTAATACTTTCAGGAGATAACAAAACAAGAGAGCAATATAAGTGTCTCCA 1650
 |||||
 Qy 488 AsnGluGluLysHisProIleIleThrSerSerLeuGlnAspGlyIleThrThrThr 507
 |||||
 Db 1651 AATGAAGAGAAACAACCAATTTACTAGCAGCTTATTACAAGATGCT-----ACTACT 1704
 |||||
 Qy 508 GlnAlaGluIleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis 527
 |||||
 Db 1705 CAAGCAGAAATTTCTACTCACTCACTATTTCAACTTCCCTACTCGAGTGTCTCACTCAT 1764
 |||||
 Qy 528 HisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrVal 547
 |||||
 Db 1765 CATGCTCAACACTTCTCTCTCTGTTGTTTCAATTCACAATGAGAGTACTACTACTGTT 1824
 |||||
 Qy 548 AspHisIleGluAsnAsnAlaLysLysGlnArgAsn---AspMetHisLysPheSerPro 566
 |||||
 Db 1825 GATCATATTGAAAAACAACGCAAAAAGCCAAAGAAATCATGCATGCACAAAGTTTCTCCA 1884
 |||||
 Qy 567 SerSerIleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGly 586
 |||||
 Db 1885 AGTAGTATTCTTTCATCTGTTGAAATGGAGCCAAAGCTAGAGAATCAACAAACAAGGG 1944
 |||||
 Qy 587 PheThrAsnProLeuMetAlaIleTyrAlaMetGlyAspPheGlyArgPheAspProHis 606
 |||||
 Db 1945 TTTACTAATCTTTAATGGCAGCTAGCGATGGAGATTTTGGAGGTTTGAATCTCAT 2004
 |||||
 Qy 607 AspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeu 626
 |||||
 Db 2005 GATCAACAATGACCCGGAATTTTCATGGAAATAATGGTGTATCTCTTACTTTAGGACTT 2064
 |||||
 Qy 627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
 |||||
 Db 2065 CTTCTCTTCTGAAACCTTAGCCATGCCAGTGAGCCAAACAAATATCTTCTTAATAGTTG 2124
 |||||
 Qy 647 GlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGluAsnIleAspPhe 666
 |||||
 Db 2125 GGAAGTAGCGCTGGAATAGGAAGTCAATACATAGATGGGATATGAAACATTTGATTT 2184
 |||||
 Qy 667 GlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheValThrGlyAsnLeu 686
 |||||
 Db 2185 CAAAGTGGGAATTAAGCGATTTTCCAACTCAACTATTGCCAGATTTTGTACAGGTAATCTT 2244
 |||||
 Qy 687 GlyThr 688
 |||||
 Db 2245 GGAACA 2250
 |||||
 RESULT 4
 AF406698
 LOCUS
 DEFINITION Solanum tuberosum BEL1-related homeotic protein 11 (Bell1) mRNA,
 partial cds.
 ACCESSION AF406698
 VERSION
 KEYWORDS AF406698.1 GI:22652116
 SOURCE Solanum tuberosum (potato)
 ORGANISM
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 1898)
 AUTHORS Chen,H., Robin,F.M., Prat,S. and Hannapel,D.J.
 TITLE Interacting transcription factors from the three-amino acid loop
 extension superclass regulate tuber formation
 JOURNAL Plant Physiol. 132 (3), 1391-1404 (2003)
 PUBMED 12857821
 REFERENCE 2 (bases 1 to 1898)
 AUTHORS Chen,H., Robin,F.M. and Hannapel,D.J.
 TITLE A KNOX protein of potato interacts with several members of the TALE

Db 1450 TTAGGGCTTCACCAAAATCTGAAGAAC-----CAACACAATTACATT 1491
Qy 643 SerAsnAspLeuGlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGlu 662
Db 1492 GGA-----TTGGAATCATTAACAATCAACCTACACATCATCAAAATATAGCTATGAA 1545
Qy 663 AsnIleAspPheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
Db 1546 AACATTGATTTCAGAGTGA--AAGCGATACGCCACTCAACTATTACAAGATTTTGT 1602
Qy 683 Thr 683
Db 1603 TCT 1605
RESULT 5
LOCUS AX506642 2043 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1337 from Patent WO0216655.
ACCESSION AX506642
VERSION AX506642.1 GI:23387879
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Harper, J.F., Krens, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 1337 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
FEATURES
source
1..2043
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 3,11e-74 Length: 2043
Score: 1150.50 Matches: 306
Percent Similarity: 51.81% Conservative: 81
Best Local Similarity: 40.96% Mismatches: 201
Query Match: 32.29% Indels: 159
DB: 6 Gaps: 28
US-10-624-201A-2 (1-688) x AX506642 (1-2043)
Qy 22 HisGlyAsn-----SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
Db 16 CACGGAACCCACCGAGATCTCTGCCGATCCGACGGTGTCTTCAACGTTGATCCTC 75
Qy 36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGln 55
Db 76 ATGAATCCAACTACTTACGTTTCAG--TACACCAACAAGACACGACTCGAACAAC 132
Qy 56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
Db 133 -----AACAAACAGCAACATAGCAACAACAACAACAACAACAACAACAAC 180
Qy 75 -----GlnHisAlaProLeu-----GlnGlnGln 82
Db 181 AACAAACAAGTAGTTTTCGTTTCTTCGATCCACGCGCGGACCAACGCGGAGCCAG 240
Qy 83 HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97
Db 241 CAGTTCGTCGGAATACCACTCTCAGGTACGAGAGCTGCTTCCATTACAGCGCGCGACAAC 300
Qy 98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGln 116
Db 301 ATCTCCGTACTTACCGGTTATTCCTCCGCGCGGTGCAGTACAGTCTCTTACGGTAGCCACAA 360

Qy 117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
Db 361 GTG----- 363
Qy 137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154
Db 364 -----GATCCCACTCACCAAGCAAGCGCG 387
Qy 155 -----GlnGlnGlnGlnGlnGlyLeuSerLeuSerSerProGlnLeuGlnGln 172
Db 388 TGTGAGACGCCGCGCGAGCGCTCTCTTTAACTCTCTGCTCTCAACAGCAGCAG 447
Qy 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerProArgThr 187
Db 448 CAACAGCAACATCATCAACAACACCGCTATTTCACGTCGATTCGGGTCCGGACATGGA 507
Qy 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202
Db 508 GAAGATATCCGGGTCCGGTCTGCTCTACAGAGATCGGGGTAAACAAACGGTATAGCTAAT 567
Qy 203 MetValIleGlySerLysTyrLeuLysAlaAlaGlnGlnLeuLeuAspGluValValAsn 222
Db 568 CTTGTT---AGCTCCAAGTACTTTGAAGGAGCAGCAAGAGCTTCTTACGAAGTAGTCAAC 624
Qy 223 IleValIleGlySerLysIleLysGlyAspAspGln-----LysLysAspAsnSer 238
Db 625 GCTGATTCCGATGACATGAACGCTAAATCCCACTATTCTCATCAAAAAGGGTAGTTGC 684
Qy 239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
Db 685 GGAATATGATAAACCTGTCTGGAGAAATCATCGCGCGCGCTGGAGAGAGAGGTTCGGGTGC 744
Qy 258 GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrAlaGlnArg 277
Db 745 GGAGCAGAGACGCGCGGAAACGT-----CCGGTGGAGCTAGGCACGCGCAGAGAGA 795
Qy 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluValGluGlnArg 297
Db 796 CAAGAAATACAGATGAAGAAAGCAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 855
Qy 298 TyrArgGlnTyrHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
Db 856 TATAGACAGTACCACGACGAGATGTCAGATGCTGATCTCTTCGTTCGACGAAGCGCAGGG 915
Qy 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db 916 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTTGAAAACCATATATCAACAGTTCCTG 975
Qy 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 976 TGCTTGAAGAGCGGATCGCTGTCAGATAAAAGCGCCCAACAGAGTCTTTGGGAGGAA 1035
Qy 358 GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHis 375
Db 1036 GATTACGTGCTGCTGCTGCTGGAGGTTTGGAGGCTCGAGGCTCAAGTTCTGGTGGACCAAC 1095
Qy 376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla 392
Db 1096 TTGAGACAGCAAGAGAGCTCTTCAACACTGGGGAATGATTCACATCTCTCAATATGCT 1155
Qy 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db 1156 TGGAGACCTCAACGCTGCTCCAGAACGAGCGCTCTCAGTTCTCCGTGCTGGCTCTTC 1215
Qy 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 1216 GAACACTCTTCTCATCATCCCTCAGGATTCGGAAGCAGCAGTGTAGTAGCAACA 1275
Qy 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db 1276 GGACTCACTCGTAGCCAGGTGTGAACTGGTTTATAAACCGGAGAGTTCGGTTATGAAA 1335

US-10-624-201A-2 (1-688) x BT008422 (1-2043)

QY	22	HisGlyAsn-----SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu	35	QY	338	CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu	357
Db	16	CACGGAAACCCACCGGAGATCTCTGCGGATCCGACGGTGGTCTTCAACAGTTGATCCTC	75	Db	976	TGCTTGAAGAGCGCATCGCTGGTCAGATAAAGCGGCCCAACAAGAGTCTTGGGGAGGAA	1035
QY	36	MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln	55	QY	358	GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHisHis	375
Db	76	ATGAATCCAACACTTACTACGTTTCAG---TACACCCAACAGACAAACGACTCGAAACAAC	132	Db	1036	GATTCACTGCTGGTGTGGAGGTTTGGAGGTCGAGGTCGAGGTCGAGTTCGGTGGACCAACC	1095
QY	56	LeuLeuPheLeuAsnSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle---	74	QY	376	LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla	392
Db	133	-----AACCAACAGCAACAATAGCAACAACAACAACAACAACAACAACAACA	180	Db	1096	TTGAGACACAAGAGCTCTTCAACAACCTGGAGATGATTCACATCTCTCCAAATATGCT	1155
QY	75	-----GlnHisAlaProLeu-----GlnGlnGln	82	QY	393	TyrArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe	412
Db	181	AACAACAACAGTAGTTTCGTTTCTCGATTCTCCACCGCGCGACCAACCGGAGCCAG	240	Db	1156	TGGAGACCTCAACGGTCTCCAGAACGAGCGCTCAGTTCCTCGTGTCTGCTCTTC	1215
QY	83	HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln	97	QY	413	GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr	432
Db	241	CAGTTCGTGCGAATACCACTCTCAGGTACAGAGCTGCTTCCATTACAGCGCGCGACAAC	300	Db	1216	GAACACTTTCTTCATCCATACCTTAAGGATTCGGACAAGCACATGCTAGCTAAGCAACA	1275
QY	98	IleAsn---HisHisGlyLeuLeuGlnAtqMetTrpAsnAsnGlnAspGlnSerGlnGln	116	QY	433	GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys	452
Db	301	ATCTCGTACTTCACGGTTATCTCCCGCGGTGACGTACAGTCTCTACGGTAGCCACCAA	360	Db	1276	GGACTCACTCTGAGCCAGGTGTCGAACCTGTTTATAAACCGAGAGTTCGGTTATGGAAA	1335
QY	117	ValIleValProSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp	136	QY	453	ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGlnGlnAsnSerThrAsn	472
Db	361	GTG-----	363	Db	1336	CCAATGGTGGAGAGATGTACATGGAGGAAATGAGGAGCAGCAGCAAGAACATG---	1389
QY	137	LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln-----	154	QY	473	ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluLysHis	492
Db	364	-----GATCCCACTCACCAAGCAAGCGCG	387	Db	1390	-----GGATCCATGGAAAGACTCTCTTGGATCAAAAGCAAGCAAGAT-----	1431
QY	155	-----GlnGlnGlnGlnGlnGlyLeuSerLeuSerProGlnLeuGlnGln	172	QY	493	ProIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSer	512
Db	388	TGTGAGACGCCACCGCGAGCAAGGCTCTCTTTAACCTCTGCTCAACAGAGCAG	447	Db	1432	-----TCTGCTTCAAAGTCA	1446
QY	173	GlnIleSerPheAsnAsn-----AsnIleSerSerSerProArgThr	187	QY	513	ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis	530
Db	448	CAACAGCAACATCATCAACAACACCAGCTTATTCAGTTCGGATTCGGGTCGGACATGGA	507	Db	1447	ACAAGTAACCAAGAAAGAGCCCAATGCGCGACACTAATATCCATATGAATCCCAATCAC	1506
QY	188	AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn	202	QY	531	AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle	550
Db	508	GAAGATATCCGGTCTCGGCTGCTGCTTACAGGATCGGGGTAAACAACCGGTATAGCTAAT	567	Db	1507	AACGGTGCCTAGAGGCGCTCCTGGAATGCAAGGAAGAGCCCAAGAGACTAAGAACACGC	1566
QY	203	MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn	222	QY	551	GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu	570
Db	568	CTTGTT---AGCTCCAAGTACTTGAAGGCAAGCAAGAGCTTCTTGACGAAGTAGTCAAC	624	Db	1567	GACGACAAATGATGCGCCAATAAT-----GCGGATTCAGCTCCAACGAGAGCTC	1620
QY	223	IleValGlyLysSerIleLysGlyAspAspGln-----LysLysAspAsnSer	238	QY	571	SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro	590
Db	625	GCTGATTCGATGACATGACGCTAAATCCCACTATTCTCATCGAAAGGGTAGTTCG	684	Db	1621	ACGATGAAGATTCTAGAAGAACGGCAAGGATAAGATCAGATGGTGGCTAC-----CCT	1674
QY	239	MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly	257	QY	591	LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis	606
Db	685	GGAAATGATAAATCTGTCGAGAATCATCGCGCGCGCTCGAGGAGAAAGGTTCCGGTGGC	744	Db	1675	TTTCATGGGTAATTCGGGCAATACCAATGGATGAGATGTCAGATTTGATGTAGTCTCA	1734
QY	258	GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg	277	QY	607	AspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLeuThrLeu	624
Db	745	GGAGCAGAACGACCGGGAACGT-----CCGGTGGAGCTAGGCACGGCAGAGAGA	795	Db	1735	GACGAGAGCTCATGGCGCAAGGTACTCAGGAAACAACANTGGCGTGTCTCTCAGCTTA	1794
QY	278	GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluValGluGlnArg	297	QY	625	GlyLeuProProSerGluAsnLeuAlaMetProValSerGln-----GlnAsnTyr	641
Db	796	CAAGAAATACAGATGAAGAAGCAAAACTTAGTAACATGCTTTCATGAGGTGGAGCAGAGA	855	Db	1795	GGTTTACCTCATTTGATAGTGTGCTCCACGACCATCAGGGTTTTCATGACAGACCCAC	1854
QY	298	TyrArgGlnTyrHisHisGlnMetGlnIleValLeuSerPheGluGlnValAlaGly	317	QY	642	LeuSerAsnAspLeuGlySerArgSerGluMetGly-----	653
Db	856	TATAGACAGTACCACCAAGCAGATGACATGCTCTCTCGTTCGAGCAAGCGCGAGGG	915	Db	1855	CATGGGATTCCTATAGGAGAGAGTGAATAATAGAGAAACAGAGGAATATGACCCGCC	1914
QY	318	IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg	337	QY	654	-----SerHis	655
Db	916	ATAGGATCAGCGAAGTCATACGTCGTAGCATTTGAAATCCATATCAAGACAGTTCGGT	975	Db	1915	ACCATCAATGGTGGTAGTCGACCACCAACCGGCACATTTCATCAGCGCAGCTGCGCGGCT	1974

QY	656	TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr	675
Db	1975	TACAATGGGTGAACATACAGAAC-----CAGAAGAGATATGGTGCT	2016
QY	676	GlnLeuLeuProAspPheVal	682
Db	2017	CAGTTATGCCGACTTCGTT	2037
RESULT 7			
LOCUS	AY096548		
DEFINITION	Arabidopsis thaliana putative homeodomain transcription factor (At2g35940) mRNA, complete cds.		
ACCESSION	AY096548		
VERSION	AY096548.1	GI:20465476	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 2074)		
AUTHORS	Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Arabidopsis Open Reading Frame (ORF) Clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2074)		
AUTHORS	Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA		
COMMENT	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.		
	The Salk, Stanford, PGECC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H., Cheuk,R., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.		
	Yamada,K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP /PGECC) contributed equally to this work as PIs.		
FEATURES	Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.		
source	Location/Qualifiers		
	1. .2074		
	/organism="Arabidopsis thaliana"		
	/molecule="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="2"		
	/clone="U18537"		
	/ecotype="Columbia"		
	/note="This clone is in pUNI 51."		
	1. .2074		
	/gene="At2g35940"		
	1. .2043		
	/gene="At2g35940"		
	/codon_start=1		
	/evidence=experimental		
	/product="putative homeodomain transcription factor"		
	/protein_id="AA20198.1"		
	/db_xref="GI:20465477"		
	/translation="MAAIFPHGNPPISAGSDGGLQTLIMPTTVQYVQQNDSDNNNNNSNNNNNTNTNTNNNSFVLDHAPQNASQQFVGIPLSGHEASITAAADNISVLHGYPVQVSLYGHQVDPDTHQAACETPRAQQGLSLTSSQQQQQQHQHQPIHVFGSGHGDIRVSGSTGVTNGIANLVSSKYLKAAQELLDEVNADSDMMNAKSLQLFSSKSGSCGNDKPVGSSGAGGEGSGGAERAGKRPVGLTAERQSIOMKKA.LSNLHEVQRYQYHQQMOWISSFQAAGISAKSYTSLAKTISRQRCLEKATAGQIKAAKSGIEEDSVSGVRFPGSRULKFDVHLHQORALQOGLMIQHPSSNARVPMVEEPERAVSLRAWLFEHLHPKPSDKHMLAKTGLTRSQVSNWFNARVRLWKPMVEEVTGMEKQAKNMGSMEKTPLDQSNEDSASKTSNOEKS PMADTYNHMNPNGDLEGVTMGSGPKRLRTSDTWPQINADPSNEKLTKMLIEBROGIRDGGVYPPFMGFGQYQMDMSRRFDVYSDQELMAQRYSGNNNGVSLTLGLPHCDLSLSTHHQGFMTHTGPIGRVKIGETEEYGPATNGSGSTTTTAHSSAAAAAYNGMNIQNKRYVAQLLPDFVA"		
	2044. .2074		
	/gene="At2g35940"		
	3'UTR		
	ORIGIN		
	Alignment Scores:		
	Pred. No.:	3,17e-74	Length: 2074
	Score:	1150.50	Matches: 306
	Percent Similarity:	51.81%	Conservative: 81
	Best Local Similarity:	40.96%	Mismatches: 201
	Query Match:	32.29%	Indels: 159
	DB:	8	Gaps: 28
	US-10-624-201a-2 (1-688) x AY096548 (1-2074)		
QY	22	HisGlyAsn-----SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu	35
Db	16	CACGGAACCCACCGGAGATCTCTCCGGATCCGCGGTGGTCTTCAACAGTTGATCCTC	75
QY	36	MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln	55
Db	76	ATGAATCCAACTACTTACGTTTCAG---TACACCCCAACAGACACGACTCGAACACAC	132
QY	56	LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAlaLeu	74
Db	133	-----AACACAGCAACCAATACCAACACACACACACACACACACACACACAC	180
QY	75	-----GlnHisAlaProLeu-----GlnGlnGln	82
Db	181	AACAAACACAGTAGTTTCGTTTTCCTCGATTCCCGCGCGCGCAACACGAGCCAG	240
QY	83	HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln	97
Db	241	CAGTTTCGTCGGAATACCACTCTCAGGTACAGAGCTGCTTCATTACAGCCGCCGACAC	300
QY	98	IleAsn---HisHisGlyLeuLeuGlnArgMetTyrAsnAsnGlnAspGlnSerGlnGln	116
Db	301	ATCTCGTACTTCAGGTTATCTCTCGCGCGTGCGTAGTACAGTCTCTACGGTAGCCACAA	360
QY	117	ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp	136
Db	361	GTG-----	363
QY	137	LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln-----	154
Db	364	-----GATCCCACTCACCAAGCGCGCG	387
QY	155	-----GlnGlnGlnGlnGlnGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln	172
Db	388	TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTAAACCTCTCGTCTCAACAGCAGCAG	447
QY	173	GlnIleSerPheAsnAsn-----AsnIleSerSerSerSerProArgThr	187

Db	448	CAACAGCAACATCATCAACACACACCGCCCTATTTCACGCTCGGATTCGGGTCCGGACATGGA	507
Qy	188	AsnAsnValThrIleArgGlyThrLeuAspGlySerSer	202
Db	508	GAAGATATCCGGTCCGGTCTCGGCTCTACAGGATCGGGGTAAACAAACGGTATAGCTAAT	567
Qy	203	MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn	222
Db	568	CTTGTT--AGCTCCAAGTACTTGAAGGCAGCACAGAGCTTCTTGACGAAGTAGTCAAC	624
Qy	223	IleValGlyLysSerIleLysGlyAspAspGln	238
Db	625	GCTGATTCGGATGATCAATGAACCGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC	684
Qy	239	MetAsnLysGluSerMetProLeuAlaSer--AspValAsnThrAsnSerSerGlyGly	257
Db	685	GGAAATGATAAACCCTGTCGGAGAAATCATCGCGCGCCTGGAGGAGAAGGTTCCGCTGCC	744
Qy	258	GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg	277
Db	745	GGAGCAGAAGCAGCGCGGAACGT-----CCGGTGGAGCTAGGCACGGCAGACAGA	795
Qy	278	GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg	297
Db	796	CAAGAAATACAGATGAAGAAAGCAAACTTTAGTAACATGCTTCATGATGAGGTGGAGCAGAGA	855
Qy	298	TyrArgGlnTyrHisHisGlnMetGlnIleValLeuSerPheGluGlnValAlaGly	317
Db	856	TATAGACAGTACCACACAGCAGATCGAATGGTGATCTCTCGTTCGAGCAAGCGGCAGGG	915
Qy	318	IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaLysSerLysGlnPheArg	337
Db	916	ATAGGATCAGCGAAGTCATACACGCTCGCTAGCATTTGAAACCATATCAAGACAGTTCCGT	975
Qy	338	CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu	357
Db	976	TGCTTGAAGAGCGCATCGCTGGTGCAGATAAAGCGGCCAACAGAGACTTTGGGGAGGAA	1035
Qy	358	GluGlyLeuGlyLysLeuGlyLeuGlySerArgLeuLysPheValAspHisHis	375
Db	1036	GATTCAGTGTCTGTGTGGAGGTTTGAGGGTTCGAGGCTCAAGTTCTGTGGACCCACC	1095
Qy	376	LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla	392
Db	1096	TTGAGACAGCAAGAGCTCTTCAACAACCTGGGAATGATTCAACATCCTTCCAATAATGCT	1155
Qy	393	TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe	412
Db	1156	TGGAGACCTCAACCTGGTCTCTCCAGAACGAGCCGTCTCAGTTCTCCGTGTGGCTCTTC	1215
Qy	413	GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr	432
Db	1216	GAACACTTCTTTCATCATACCCCTAAGATTCGAGCAAGCAGCATGCTAGCTAAGCAAA	1275
Qy	433	GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys	452
Db	1276	GGACTCACTGTCAGCCAGGTGTCGAACCTGGTTTATAAACCGAGAGTTCCGTTATGAAA	1335
Qy	453	ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn	472
Db	1336	CCAAATGGTGGAGGAGATGTACATCGAGGAAATGAAGGAGCAGGCGCAAGCAATC-	1389
Qy	473	ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis	492
Db	1390	-----GGATCCATGGAAAGACTCCTTTGGATCAACGACACAGAGAT-----	1431
Qy	493	ProfileThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSer	512
Db	1432	-----TCTGCTTCAAAAGTCA	1446
Qy	513	ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis	530

Db	1447	ACAAGTAACAAGAAAGAGCCCAATATGGCGGACACATAATTACCATATGAATCCCAATCAC	1506
Qy	531	AspPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle	550
Db	1507	AACGGTGACCTAGAGGGCGTCACTGGNATGCAAGGAAGCCCAAGAGACTAAGAACACGAC	1566
Qy	551	GluAsnAsnAlaLysIysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu	570
Db	1567	GACGAGACAATGATGACGACCAATAAAT-----GCGGATTTACGTCCCAACGAGAAGCTC	1620
Qy	571	SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro	590
Db	1621	ACCATGAAGATTCTAGAAGACCGCAAGGATTAAGATCAGATCGTGCTCTAC-----CCT	1674
Qy	591	LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis	606
Db	1675	TTCATGGGTAAATTCGGGCAATACCAATGGATGAGATGTCAGATTTGATGTAGTCTCA	1734
Qy	607	AspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLeuThrLeu	624
Db	1735	GACCAGAGCTCATGGCGCAAGGTACTCAGGAACAACAATGGCGTGTCCCTCAGCTTA	1794
Qy	625	GlyLeuProProSerGluAsnLeuAlaMetProValSerCln-----GlnAsnTyr	641
Db	1795	GGTTTACCTCATGTGATAGCTTGTGTCACGACCATCAGGGTTTCATGCACGCCAC	1854
Qy	642	LeuSerAsnAspLeuGlySerArgSerGluMetGly-----	653
Db	1855	CATGGGATTCCTATAGGAGAGAGCTGAAATAGGAGAAACAGAGGATATGACCGGCC	1914
Qy	654	-----SerHis	655
Db	1915	ACCATCAATGGTGGTAGCTCGACCACACCGCACATTCATCAGCGGCAGCTGCGCGGCT	1974
Qy	656	TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr	675
Db	1975	TACAAATGGGATGAACATACAGAAC-----CAGAAGAGATATGTGGCT	2016
Qy	676	GlnLeuLeuProAspPheVal	682
Db	2017	CAGTTATGGCCGACTTCGTT	2037
RESULT 8			
AY072175			
LOCUS	Arabidopsis thaliana putative homeodomain transcription factor		
DEFINITION	(At2g35940) mRNA, complete cds.		
ACCESSION	AY072175.1	GI:18176172	
VERSION	FLI CDNA.		
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.		
REFERENCE	1 (bases 1 to 2572)		
AUTHORS	Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Sato, R., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.		
TITLE	Arabidopsis Full Length cDNA Clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2572)		
AUTHORS	Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Sato, R., Seki, M., Shinn, P., Southwick, A.		

TITLE Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 JOURNAL Direct Submission
 COMMENT Submitted (02-JAN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PCEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PCEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PCEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source Location/Qualifiers
 1..2572
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosomes="2"
 /clone="RAFL09-25-N04 (R18537)"
 /ecotype="Columbia"
 /note="This clone is in a modified pBluescript vector (FLC-1) as a BamHI/XhoI insert."
 Gene 1..2572
 /gene="At2g35940"
 5'UTR 1..205
 /gene="At2g35940"
 misc_difference 2
 /gene="At2g35940"
 /note="compared to genomic sequence"
 /replaces="c"
 206..2248
 /gene="At2g35940"
 /codon_start=1
 /evidence=experimental
 /product="putative homeodomain transcription factor"
 /protein_id="AAL59997.1"
 /db_xref="GI:18176173"
 /translation="MAAYPHGNPPETISAGSGGLQTLILNPPTVVOYTQDNDNNNNNSNNNNNTNTNNNSSFVFLDHPQNAQQFVGIPLSGHEAASITAAADNISVLHGPPRVQYSLYGHQVDPHQACETPRAQGLSLTSSQQQQQHHQHOPIHYVFGSGHEDIRVSGSGVTGQIANLVSKYLKAAQELLDEVNADSDMNKASQFSSKKGCGNDPQVSGSAGGEGGGAAGKRPVELGTAEQRIQMKKAKLSNMLHEVQRYOHOQOMVIVSFEQAGIGSAKSVSLAKTISRQFRLKEAIAQIKAAKSLGEEDSVSGVRFEGRSLKFDVHLRQORALQOLCMIQHPNSNAWRPQGLPERAVSLRAWLEFHLPHLPFKDSDKMLAKOTGLTRSVSNWFINARVRLWKPVVEEITMEENKQAKMSEKPTLDQSNEDSASKSTSNQKSPMDATNHPNPNHGDLEGVTGMOGSPRLSDQELMAQRYSNNNVSLTGLPHCDLSLSTHQGFQTHHGIPIGRRVKIGETEYGPATINGGSSSTTAHSSAAAAAYNGMNIQNKRYVAQLLPDFVA"
 3'UTR 2249..2572
 /gene="At2g35940"
 misc_difference 2417..2418
 /gene="At2g35940"
 /note="compared to genomic sequence"
 /replaces="ttc"
 misc_difference 2556
 /gene="At2g35940"
 /note="not present in genomic sequence"

ORIGIN

Alignment Scores:
 Pred. No.: 4,19e-74 Length: 2572
 Score: 1150.50 Matches: 306
 Percent Similarity: 51.81% Conservative: 81
 Best Local Similarity: 40.96% Mismatches: 201
 Query Match: 32.29% Indels: 159
 DB: 8 Gaps: 28
 US-10-624-201a-2 (1-688) x AY072175 (1-2572)
 QY 22 HisGlyAsn-----SerAsnAsnAsnAsnileGlnThrLeuTyrlau 35
 |||||
 Db 221 CACGGAACCCACCGAGATCTCCGCGATCCGAGGGTGTCTTCAAACGTTGATCCTC 280
 QY 36 MetAsnProAsnAsnTyrlMetGlnGlyTyrlThrSerAspThrGlnGlnGlnGln 55
 |||||
 Db 281 ATGAATCCAACTACTTACGTTTCAG---TACACCCCAACAGCAACGACTCGAACACAA 337
 QY 56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCyHsialaAsnile--- 74
 |||||
 Db 338 -----AACACAGCAACAATAGCAACAACAACAACAACAACAACAACAACA 385
 QY 75 -----GlnHisAlaProLeu-----GlnGlnGln 82
 |||||
 Db 386 AACAAACAACAGTAGTTTGGTTTTCCTCGATTCCCGCGCGCCGAGCAAAACGAGCCAG 445
 QY 83 HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97
 |||||
 Db 446 CAGTTTCGTCGAATACCACTCTCAGGTACGAAAGCTGTTCCATTACAGCCCGCACAA 505
 QY 98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnGlnGln 116
 |||||
 Db 506 ATCTCGGTACTTCAGGGTTATCTCTCGCGCGGTGAGTACAGTCTCTACGGTAGCCACCA 565
 QY 117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyleThrThrAsp 136
 |||||
 Db 566 GTG----- 568
 QY 137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154
 |||||
 Db 569 -----GATCCCACTCACCAGCAAGCCGG 592
 QY 155 -----GlnGlnGlnGlnGlnGlyLeuSerLeuSerProGlnLeuGlnGln 172
 |||||
 Db 593 TGTGAGACGCCACGCGCGAGCAAGCCCTCTTTAAACCTCTCGTCTCAACAGCAGCAG 652
 QY 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerSerProArgThr 187
 |||||
 Db 653 CAACAGCAACATCATCAACAACACAGCCGCTATTACGTCGGATTCCGGTCCGAGCATGGA 712
 QY 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202
 |||||
 Db 713 GAAGATATCCGGGTGGGTCTGCTGCTCAGAGATCGGGGTAAACAACGGTATAGCTAAT 772
 QY 203 MetValLeuGlySerLysTyrlLeuLysAlaAlaGlnGlnLeuLeuAspGluValValAsn 222
 |||||
 Db 773 CTGTGT---AGCTCCAAGTACTTGAAGCAGCACAAGAGCTTCTTGGACGAAGTAGTCAAC 829
 QY 223 IleValGlyLysSerIleLysGlyAspAspGln-----LysLysAspAsnSer 238
 |||||
 Db 830 GCTGATTCGATGATGATCAACGCTAAATCCCACTATTCTTCATCGAAAAAGGTTAGTTC 889
 QY 239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
 |||||
 Db 890 GGAATATGAACCTGTTCGAGAAATCATCGCGCGCGCTGGAGGAGAGAGTTCCGGTGGC 949
 QY 258 GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
 |||||
 Db 950 GGAGCAGAAGCAGCCGGAAACGT-----CCGGTGGAGCTAGGCACGCGCAGAGAGA 1000
 QY 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGlnArg 297
 |||||
 Db 1001 CAAGAAATACAGATGAAGAAAGCAAACTTAGTAACTGCTTTCATGAGGTGGAGCAGAGA 1060

```
Qy 298 TyrArgGlnTyrHisHisGlnMetGlnIleValLeuSerPheGluGlnValAlaGly 317
Db 1061 TATAGACAGTACCACGACGATGCAGATGCTGATCTCTTCGTTCCGAGCAAGCGCAGGG 1120
Qy 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaLysLysGlnPheArg 337
Db 1121 ATAGGATCAGCGAAGTCATACACGTCGTCAGCATTTGAAACCAACCATATCAAGACAGTTCCTCGT 1180
Qy 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 1181 TGCTTGAAGAGCGATCGCTGTCAGATTAAGAGCGGCCCAACAGAGTCTTTGGGAGGAA 1240
Qy 358 GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
Db 1241 GATTTCAGTGTCTGGTGTGGAGGTTTTCAGGGGTCGAGGCTCAAGTTTCGTGGACCAACCAC 1300
Qy 376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla 392
Db 1301 TTGAGACAGCAAGAGCTCTTCAACAACCTGGGAATGATTCAACATCCTTCCAAATAATGCT 1360
Qy 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db 1361 TGGAGACTTCAACGTGCTCTCCAGAACGAGCCGCTTCAGTTCTCCGTGCTTGGCTCTTC 1420
Qy 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 1421 GAACACTTCTTCATCATACCCTAAGGATTTCGGACAAAGCACATGCTAGCTAAGCAACA 1480
Qy 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db 1481 GGACTCACTCGTACCCAGGTGTCCAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1540
Qy 453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
Db 1541 CCAATGTTGGAGGAGTGTATCATCGAGGAATGAGGAGGAGCGCAAGAACATG----- 1594
Qy 473 ThrSerGlyAspAsnLysAsnLysGluThrAsnLysSerAlaProAsnGluGluLysHis 492
Db 1595 -----GGATCCATCGAAAGAGACTCTTTGGATCAAAGCAACGAAGAT----- 1636
Qy 493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSer 512
Db 1637 -----TCTGTTCAAAGTCA 1651
Qy 513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis 530
Db 1652 ACAAGTAACCAAGAAAGACCCCAATCGCGACACTAATTACCATATGATCCCAATCAC 1711
Qy 531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
Db 1712 AACGGTGCACCTAGAAGCGCTCACTCGGAATGCAAGGAAGCCCAAGAGACTAAGAACCCAGC 1771
Qy 551 GluAsnAsnAlaLysLysGlnArgAsnAspMethHisLysPheSerProSerSerIleLeu 570
Db 1772 GACGAGCAATGATGCGCAATAAAT-----CCGGATTTCAGCTCCCAAGCAAGAACTC 1825
Qy 571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db 1826 ACGATGAAGATTCTAGAAGAACGCGCAAGGNTAGATCATGATGTTGGCTAC-----CCT 1879
Qy 591 LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db 1880 TTCATGGGTATTTCCGGGCAATACCAATGGATGATGATGTCAGATTTGATGTAGTCTCA 1939
Qy 607 AspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLeuThrLeu 624
Db 1940 GACCAGGAGCTCATGGCGCAAGAGTACTCAGGAACAACCAATGCGCGTTCCTCAGCTTA 1999
Qy 625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln-----GlnAsnTyr 641
Db 2000 GGTTTACTCTATTGTGATAGTCTGCTCCAGCACCATCAGGGTTTCATGTCGACACCCAC 2059
```

```
Qy 642 LeuSerAsnAspLeuGlySerArgSerGluMetGly----- 653
Db 2060 CATGGGATTCCTATAGGAGAGAGTGAATAATAGAGAAACAGAGGAATATGGACCCGCC 2119
Qy 654 -----SerHis 655
Db 2120 ACCATCAATGGTGTAGTCGACCAACACCGCACTTCATCAGCGGAGCTGCGCGGCT 2179
Qy 656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675
Db 2180 TACAATGGATGAACATACAGAAC-----CAGAAGAGATATGTGGCT 2221
Qy 676 GlnLeuLeuProAspPheVal 682
Db 2222 CAGTTATGCGCGACTTCGTT 2242

RESULT 9
AY099854 2578 bp mRNA linear PLN 06-MAY-2002
LOCUS Arabidopsis thaliana putative homeodomain transcription factor
DEFINITION (At2g35940) mRNA, complete cds.
ACCESSION AY099854.1 GI:20466775
VERSION AY099854
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 2578)
AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
Palmer,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
TITLE Submitted (24-APR-2002) DNA Sequencing and Technology Center,
JOURNAL Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PCEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs (RAFL cDNA : 'RIKEN
Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
source
1. 2578
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAFL09-94-K12 (R25364)"
/ecotype="Columbia"
/notes="This clone is in pBluescript"
1. 2578
/gene="At2g35940"
/notes="synonym: FLIP19.15"
198..2240
/gene="At2g35940"
/codon_start=1

gene
CDS
```

```
/product="putative homeodomain transcription factor"  
/protein_id="AA20705.1"  
/db_xref="GI:20466776"  
/translation="MAAYFHNPPEISAGSDGLQTLILNPTTVYVYQQNDSDNN  
NNSNNNNNTNTNNNSPFLDSHAPQPNASQFVGIPLSGEAAASITAADNIS  
VLHGYPYQVYLSGHQVDPLTHQQACETPRAGQSLTSLSSQQQQQHQHQPI  
HVGFGSGHEDIRVSGSGTGVNGIANLVSKYLAQAQELLDEVNADSDMMNAKS  
OLPSSKSGCNDKPVGESSAGAGGSGGGAEEAKRPVELGTAEQRIOMKKAKLS  
NMLHEVQRYOYHQOMQVLSFEEQAGISAKSVTSLAKTISRQPRCLKEATAGQ  
IKXAKSLGEEDSVSGRPEGRSLKFPVDHLHQQRALQQLGMIQHPSNNAWRPQRL  
PERAVSLVRLUFEHLFHPYKDSDKMLAKQTLGTRSQVSNWFINARVRLKPMVBE  
MYMEKEQAKNMSMEKTPLOQSDNEDSASKTSNOEKSFMADTNYHNNPNHGLEG  
VTGMQSGRLKRLTADPMQPINADPSNNEKLTMLLEBRQGRIRSDGYPFNGNFOY  
QNDMSRFQVSDQELMAQRYSGNNNGVSLTLGLPHCDLSLSTHOGEMQTHGPIG  
RRVKIGTEYGPATINGSSTTAHSSAAAAAYNGMIQNKRYVAQLLPDFA"
```

ORIGIN

Alignment Scores:

Pred. No.:	4,21e-74	Length:	2578
Score:	1150.50	Matches:	306
Percent Similarity:	51.81%	Conservative:	81
Best Local Similarity:	40.96%	Mismatches:	201
Query Match:	32.29%	Indels:	159
DB:	8	Gaps:	28

US-10-624-201A-2 (1-688) x AY099854 (1-2578)

QY	22	HisGlyAsn-----	SerAsnAsnAsnIleGlnThrLeuTyrLeu 35
DB	213	CACGGAACCCACCGGAGATCTCTGCGGATCCGACGGTGTCTTCAAACTTTGATCCTC 272	
QY	36	MetAsnProAsnAsnTyrMetGlnGlyTyrThrSerAspThrGlnGlnGlnGln 55	
DB	273	ATGAATCCAACTACTTACGTTTCA-----TACACCAACAAGAACACGACTCGAACACAAC 329	
QY	56	LeuLeuPheLeuAsnSerProAlaAlaSerAsnAlaLeuCysHisAlaHisIle--- 74	
DB	330	-----AACACAGCAACAATAGCAACAACAACAACAACAACAACAACAACAAC 377	
QY	75	-----GlnHisAlaProLeu-----GlnGlnGln 82	
DB	378	AACAACAACAGTAGTTTCGTTTTCTCGATTCCACGCGCGCGAGCCAAACGCGAGCCAG 437	
QY	83	HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97	
DB	438	CAGTTTCGGAATACACTCTCAGTGCACGAGCTGCTTCCATTACAGCGCGCGCAAC 497	
QY	98	IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116	
DB	498	ATCTCGGTACTTCACGGTTATCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCA 557	
QY	117	ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrAsp 136	
DB	558	GTG----- 560	
QY	137	LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154	
DB	561	-----GATCCCACTCACAGCAAGCGCGG 584	
QY	155	-----GlnGlnGlnGlnGlnGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172	
DB	585	TGTGAGAGCCACGCGCGAGCAAGCCCTCTCTTTAACCTCTCTCTCAACAGCAGCAG 644	
QY	173	GlnIleSerPheAsnAsn-----AsnIleSerSerSerSerProArgThr 187	
DB	645	CAACAGCAACATCATCAACAACACCGCTATTTCACGTGGATTTCGGGTCCGACATGGA 704	
QY	188	AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202	
DB	705	GAAGATATCGGGTCGGGTCTGCTCTACAGGATCGGGGTAAACAACCGGTATAGCTAAT 764	
QY	203	MetValLeuGlySerIlystYrLeuLeuAlaGlnGlnLeuLeuAspGluValValAsn 222	

DB	765	CTTGT---AGTCCCAAGTACTTGAAGCGAGCACAAAGAGCTTCTTGACGAAGTAGTCAAC 821	
QY	223	IleValGlyIlySerIleIysGlyAspAspGln-----LysIysAspAsnSer 238	
DB	822	GCTGATTCGATGCATGAACCGCTAAATCCCACTATTCTCATCGAAAAAGGGTAGTTGC 881	
QY	239	MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257	
DB	882	GGAAATGATAACCTGTCGAGAAATCATCGCGCGCGCTGGAGGAGAAAGTTCGGTGGC 941	
QY	258	GlyGluSerSerSerArgGlnIlySerGlnValAlaValGluLeuThrThrAlaGlnArg 277	
DB	942	CGAGCAAGAACCGCGGAAACGT-----CCGGTGGAGCTAGGCACGCGCAGAGAGA 992	
QY	278	GlnGluLeuGluMetIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 297	
DB	993	CAAGAAATACAGTGAAGAAAGCAAACTTAGTAACTGCTTCATGAGTGGAGCAGAGA 1052	
QY	298	TyrArgGlnTyrHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317	
DB	1053	TATAGACAGTACCACGACGATGCAGATGGTGTATCTCTTCGTCGAGCAACGCGCAGG 1112	
QY	318	IleGlySerAlaIlySerTyrThrGlnLeuAlaLeuHisAlaIleSerIysGlnPheArg 337	
DB	1113	ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTTGAAACCATATCAACAGAGTTCCT 1172	
QY	338	CysLeuLysAspAlaIleAlaGluGlnValIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 357	
DB	1173	TGCTTGAAGAGCGCATCGCTGTCAGATAAAGCGGCCAACAAAGAGTCTTGGGGAGGAA 1232	
QY	358	GluGlyLeuGly-----GlyIlyIleGluGlyIlySerArgLeuLysPheValAspHis 375	
DB	1233	GATTCACTGCTGCTGCTGGAGGTTTGGAGGCTCGAGGCTCAAGTTCGTCGACCAACC 1292	
QY	376	LeuArgGlnGlnArgAlaLeuGlnIleGlyMetMetGln-----ProAsnAla 392	
DB	1293	TTGAGACAGCAAGAGCTCTTCAACAACTGGGAATGATTTCAACATCTCTTCCATATGCT 1352	
QY	393	TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412	
DB	1353	TGAGACCTTCAACGTGCTCCAGAACGAGCGCTCAGTTCCTCGTCTCGGTCTTC 1412	
QY	413	GluHisPheLeuHisProTyrProIlySerAspSerAspLysIleMetLeuAlaLysGlnThr 432	
DB	1413	GAACACTTCTTTCATCATACCTAAGGATTCGACAAACACATGCTAGTAAAGCAACA 1472	
QY	433	GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452	
DB	1473	GGACTCACTCGTAGCCAGGTGTGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGA 1532	
QY	453	ProMetValGluGluMetTyrLeuGluValLysAsnGlnGlnGlnAsnSerThrAsn 472	
DB	1533	CCAATGTCGAGAGATGATACATGGAGGAATGAAGGACGAGCAAGAACATG----- 1586	
QY	473	ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492	
DB	1587	-----GGATCCATGGAAAGACTCCTTTGGATCAAAAGCAACGAAGT----- 1628	
QY	493	ProIleIleThrSerSerLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512	
DB	1629	-----TCTGCTTCAAAAGTCA 1643	
QY	513	ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis 530	
DB	1644	ACAAGTAACAAGAAAGAGCCCAATGGCGGACACTAATACCATATGATTAATCCCAATCAC 1703	
QY	531	AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550	
DB	1704	AACGGTGACCTAGAGCGCTCACTGGAATGCAAGGAAGCCCCCAAGAGACTTAAGAACACGC 1763	
QY	551	GluAsnAsnAlaLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570	
DB	1764	GACGAGACAATGATGCAGCCAAATAAT-----GCGGATTTTCACTCCCAACGAGAGCTC 1817	

QY	571	SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro	590
Db	1818	ACGATGAAGATTCTAGAGAAGACGCAAGGATAAGATCATAGTGGCTAC	1871
QY	591	LeuMet	606
Db	1872	TTTCATGGGTAATTTCGGGCAATACCAATCGATGATGTCAGATTGTAGTCTCA	1931
QY	607	AspGlnGlnMetThrAlaAsnPheHis	624
Db	1932	GACAGAGAGCTATGGCGCAAGAGTACTCAGGAACCAACATGGCTGCTCCCTCACGTTA	1991
QY	625	GlyLeuProSerGluAsnLeuAlaMetProValSerGln	641
Db	1992	GGTTTACCTCATGTGTAGTACTGTGCTCCAGCACCATCAGGTTTCATGCAGACCCAC	2051
QY	642	LeuSerAsnAspLeuGlySerArgSerGluMetGly	653
Db	2052	CATGGGATTCCTATAGGGAAGAGTGAGTGAATAAGGAGAAACAGAGGAATATGACCCGCC	2111
QY	654		655
Db	2112	ACCATCAATGGTGTAGTCGACCACCAACCGCACATTCATCAGCGGCGAGCTGCGGGCT	2171
QY	656	TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr	675
Db	2172	TACAATGGGATGAACATACAGAAC	2213
QY	676	GlnLeuLeuProAspPheVal	682
Db	2214	CAGTTATGCGCGACTTCGTT	2234
RESULT 10			
AF353094			
LOCUS	AF353094	2385 bp	linear
DEFINITION	Arabidopsis thaliana BEL1-like homeodomain 1 (BLH1) mRNA, complete cds.		
ACCESSION	AF353094		
VERSION	AF353094.1		
KEYWORDS	GI:13877516		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 2385)		
AUTHORS	Pidkowich, M.S., Samach, A., Modrusan, Z. and Haughn, G.W.		
JOURNAL	Unpublished		
TITLE	A family of BEL1-like homeodomain (BLH) proteins in Arabidopsis thaliana		
FEATURES	source		
gene			
CDS			

/db_xref="GI:13877516"

/translation="MAAYFHGNPPEISAGSDGLQTLILMNPITYVQYVYQTDNDNNNNNSNNNTNTNTNNSSFFVLDHSHAPQPNASQFVGIPLSGHEAASTITADNISVLHGYPVQYSLYSHOVDPTHQAACTTPRAOGLSLTSLSSQOQOQHOHQHOPIHYVFGSGCEDIRVSGSGTGVNGIANLVSKYLKAAQELLDEVDNADSDDNKASQLFSSKGGCGNDKPKVGSAGAGSGGGAAGAKRPVELGTARQELQMKKAKULSNLHVEQRYRQHQQOMVSISSFEAAGISAKSYTSLAKTISIRQFRCLEKAIAGQIKAAKNSLGEEDSVSGVGRFEGSRLLKFDVPHLRQOALQOGLMIQHPSSNNAWRPQGLPERAVSVLRWLIFELHPYKPDOSDKMLAKOTGLTRSOVSNWFINARVLRKMPWLEBVTMEEMKEAKNGMSMEKTPLDOSNEDSASKTSNOEKS PMADTNTYHNNHNDLEBVTMOCPCRLRTSDETMMQPINADPSSNEKLTMLKLEERQGISDGGYFPMGNFGQYQNDMSRFDVSDOELMAQRYSGNNNGVSLTLGLPHCDLSLSDTHQGYPMOHTGPIPGRRVKIGETBEYGEATTINGSSITTAHSSAAAAAAYNGMNIQNKRYVYVQQLLPDFA"

Alignment Scores:	5 31e-74	Length:	2385
Pred. No.:	1148.50	Matches:	306
Score:	51.67%	Conservative:	80
Percent Similarity:	40.96%	Mismatches:	202
Best Local Similarity:	32.23%	Indels:	159
Query Match:	8	Gaps:	28
DB:			

US-10-624-201A-2 (1-688) x AF353094 (1-2385)

QY 22 HisGlyAsn-----SerAsnAsnAsnIleGlnThrLeuTyrLeu 35

Db 194 CACGGAACCCACCGGAGATCTCTCGCGGATCCGACGGTCTTCAACAGTTGATCCTC 253

QY 36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln 55

Db 254 ATGAATCCAACTACTTACGTTTCAG--TACACCCAAACAAGACCAACGACTCGAACAAAC 310

QY 56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCyHisAlaAsnIle-- 74

Db 311 -----AACACAGCAACAATAGCAACAACAACAACAACAACAACAACAACAACA 358

QY 75 -----GlnHisAlaProLeu-----GlnGlnGln 82

Db 359 AACACAACAGTAGTTTCTGTTTCTCGATCCACGCGCGCGACCAACCGGAGCGAG 418

QY 83 HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97

Db 419 CAGTTCGTCGGAAATACCACTCTCAGTCACGAAGCTGCTTCCATACAGCCGCGGCAAC 478

QY 98 IleAsn---HisHisGlyLeuLeuGlnAbgMetTyrAsnAsnGlnAspGlnSerGlnGln 116

Db 479 ATCTCGTACTTCACGGTTATCTCTCGCGGCTGCGAGTACAGTCTCTACGGTAGCCCAA 538

QY 117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136

Db 539 GTG----- 541

QY 137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154

Db 542 -----GATCCCACTCACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 565

QY 155 -----GlnGlnGlnGlnGlnGlnGlyLeuSerLeuSerProGlnLeuGlnGln 172

Db 566 TGTGAGACGCCGCGCGCGAGCAAGGCTCTCTTTAACCTCTCGTCTCAACAGCAGCAG 625

QY 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerSerProArgThr 187

Db 626 CAACAGCAACATCATCAACAACACAGCCCTATTTCAGTCGGATTCGGGTCGGACATGGA 685

QY 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202

Db 686 GAAGATATCCGGTCTGGCTCTGAGGATCGGGGTAAACAACCGGTATAGTAAAT 745

QY 203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGlnLeuLeuAspGluValValAsn 222

Db 746 CTTGTT---AGCTCCAGTACTTGAAGGCGACAAAGAGCTTTTTCAGCAAGTAGTCAAC 802

/protein_id="AA03626.1"
 /db_xref="GI:22652125"
 /translation="QGLSLSSQQPFGNFTAAARELVSSPSGSASAGIOOQOQOQ
 QSISSVPLSSKYMKAAQELLDVNVNKGSMKSTNSDVTNVNNDVKSKMGMDDGOLD
 GVGDADKDGATSTELSTGERQEIOMKKAKLVNMLDSEVEORYHYHOMQSVIHWLEAA
 GIGSAKTYTALAQTTISKQPRCKDAIIGQIRASQTLGEEDSLGKIGSRUKFVDN
 QLRQOALQOQGMQIHNMRPQGLPERAVSVLRWLFEHLPHYPKDSKMLAKOT
 GLTRQSVNFWINARVLKPMVEYMLLEEIKHEQNGLGQEKTSKLQONEDSTTSR
 SIATQKSPGSDSQNKSFVSKQDNHLPHQNPSPMPDQVORHPFHTPTGMTIRNQSAGFN
 LIGSPITESINITQSPKPKRNNEMLHSPNSIPSINMDVKPNEQMSKMGFDROORD
 GFSLMGPMFMGFGFAYPIGETARSTQFSPAPYSTGCTVSLTLGLPHNENLSMAT
 HHSFLPIPTQNIQIGSEPNHFEGLNTPISAHSTSSVYETFNIGNRRRAAPLUPDFV
 A"

ORIGIN

Alignment Scores:
 Pred. No.: 1 55e-72 Length: 2128
 Score: 1127.50 Matches: 273
 Percent Similarity: 59.52% Conservative: 77
 Best Local Similarity: 46.43% Mismatches: 150
 Query Match: 31.64% Indels: 89
 DB: 8 Gaps: 18

US-10-624-201A-2 (1-688) x AF406702 (1-2128)

Qy 161 GlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnGlnSerPheAsnAsn----- 178
 Db 4 GGGCTTTCACCTAGCCTGCTCCTCG-----TCCAGCAGCGGGGTTTGGGAACCTTCACG 57
 Qy 179 -----AsnLeuSerSerSerSerProArgThrAsnAsnValThrIle 192
 Db 58 GCGGCGCGTGAGCTTGTTCTTCGCTTCGGGTTTCGGCTTCAGCTTCAGGATACACAA 117
 Qy 193 ArgGlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLeuValLeuValA 212
 Db 118 CAACAACAGCAACAAACAGAGTATTAGTAGTGTGCTTTAGTCTTAAGTACATGAAGGCT 177
 Qy 213 AlaGlnGlnLeuLeuAspGluValValAsnIleValGlyLyssSerIleLys----- 229
 Db 178 GCACAGAGCTACTTGATGAAGTTGTAAAT--GTTGGAANAATCAATGAAGAAGTACTAAT 234
 Qy 230 -----GlyAspAspGlnLysLysAspAsnSerMetAsnLysGluSer 243
 Db 235 AGTACTGATGTTGTTGTTAATAATGATGTCAGAAATCGAAGAAATATG----- 282
 Qy 244 MetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSerArg 263
 Db 283 -----GGCGATATGGACGGACAGTTAGACGGAGTTGGA-----GCAGAC 321
 Qy 264 GlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGlnLeuGlnMetLys 283
 Db 322 AAAGACGGAGCTCCAACAACCTGAGCTAAGTACAGGGGAGAGACAAAGAAATTCAAATGAAG 381
 Qy 284 LysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHis 303
 Db 382 AAAGCAAAACTTGTAACTGCTTGACAGGTGAGCAGAGGTATAGACATATATCATCAC 441
 Qy 304 GlnMetGlnIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSer 323
 Db 442 CAATGTCAGTCAGTGATACATTTGGTTAGAGCAGAGCTGCTGTCATTTGATCAGCAAAACA 501
 Qy 324 TyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIle 343
 Db 502 TATACAGCATTTGGCTTTGCAGACGATTTTCAGCAATTTTAGTGTCTTTAAGGACGCGATA 561
 Qy 344 AlaGlnValLysAlaThrSerLysSerLeuGlyGluGlnGluGlyLeuGlyGlyLys 363
 Db 562 ATTGGTCAAAATACGATCAGCAAGCAGCAGCTTAGGCGAAGAAAGATAGTTTGGGAGGGA 621
 Qy 364 IleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGln 383
 Db 622 ATTGAAGGTTCAAGGCTTAAATTTGTTGATAATCAGCTAAGACAGCAAGGCGCTTTTGCAA 681

Qy 384 GlnIleGlyMetMetGlnProAsnAlaTrrArgProGlnArgGlyLeuProGluArgAla 403
 Db 682 CAATTGGGAATGATCCAGCATATATCTTGAGACCTTCAGAGAGGATTCGCCCAAGAGCT 741
 Qy 404 ValSerValLeuArgAlaTrrLeuPheGluHisPheLeuHisProTyrProLysAspSer 423
 Db 742 GTTTCCTGTTCTCGCGCTTGGCTTTTGAACATTTCTCTCCATCTTATCCCAAGGATTC 801
 Qy 424 AspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrrPhe 443
 Db 802 GACAAATGATGTACAAAACAAACAGAGCTAATCTAGGAGTCAGGTGTCTGATTTGGTTC 861
 Qy 444 IleAsnAlaArgValArgLeuTrrLysProMetValGluGluMetTyrLeuGluVal 463
 Db 862 ATCAATGCTCGAGTTGCTTTTGGAAAGCCAATGGTGAAGAGATGTACTTGGAGAGATA 921
 Qy 464 LysAsnGlnGluGlnAsn-----SerThrAsnThrSerGlyAsp---Asn 477
 Db 922 AAAGAACACGAAACAGAAATGGTTGGGTCAAGAAAACAGAGCAAAATTAGGTGAACAGAAC 981
 Qy 478 LysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSer 497
 Db 982 GAAGATTCACACATCAAGATCCATGTGTACACAGACAAAGCCCTGGTTCAGATAGC 1041
 Qy 498 SerLeuLeuGlnAspGlyIleThrThrGlnAlaGluLeuSerThrSerThrIleSer 517
 Db 1042 CAAAA-CAAGAGTTTGTCTCAAAAACAGGACAATCATTTGCCTCAACACACACCTGCTTC 1100
 Qy 518 ThrSer-ProThrAlaGlyAlaSerLeuHis-----Hi 528
 Db 1101 ACCAATGCCCGATGTCCAAACGCCACTTCCATACCCCTATCGGTATGACCAATCCGTAATCA 1160
 Qy 528 sAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAs 548
 Db 1161 GTCTGTGTTTCAACTCATTCGATCACCAGAGATCGAAGATCAACACTACT----- 1215
 Qy 548 pHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSe 568
 Db 1216 -----CAAGGGAGTCCAAAGAAACCGAGGAAACACAGAGATGTTCATTTCACCAACAG 1268
 Qy 568 rIleLeuSerSerValAspMetGluAlaLysAlaArgGlu----- 581
 Db 1269 CATT-----CCATTCATCAACATGATGTAAAGCCCTAACGAGGAAACAAATGTCGATGAAGTT 1325
 Qy 582 -----SerSerAsnLysGlyPheThr-----AsnProLeu----- 591
 Db 1326 TGGTGATGATAGCAGCAGACAGATGGATTCCTCACTAATCGGAGGACCGATGAACCTCAT 1385
 Qy 592 -----MetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspProHisAspGlnG 609
 Db 1386 GGGAGGATTCGGAGCCTATCCCATTCGAGAAATTCGTCGGTTTAGC-----ACCGAGCA 1439
 Qy 609 nMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProProSe 629
 Db 1440 ATTCTCAGCACCATTATCAACACCGATGGCAGAGTTTCACTCACTCTTGGCCTACACATAA 1499
 Qy 629 rGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeu----- 642
 Db 1500 CGAAAACTCTCAATGTCTGCAACACACACAGCTTTCCTTCCATTCCAACACAAAAACAT 1559
 Qy 643 -----SerAsnAspLeuGlySer---ArgSerGluMetGlySerHis 655
 Db 1560 CCAATTTGGAAAGTGAACCAAAATCATGATTTGGTAGCTTAAACACACACCAACATCAGCTCA 1619
 Qy 655 sTyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProTh 675
 Db 1620 CTCACATCAAGCGTCTATGAACCTTCACATTCAGAAC-----AGAAAGAGGTTTCGCCGC 1676
 Qy 675 rGlnLeuLeuProAspPheVal 682
 Db 1677 ACCCTTGTACAGATTTTGT 1698

RESULT 12


```

repeat_region 22577..22612
/rpt_family="AT_rich"
repeat_region 23077..23102
/rpt_family="AT_rich"
repeat_region complement(25817..25860)
/rpt_family="AT_rich"
repeat_region complement(25894..25930)
/rpt_family="AT_rich"
26200..26382
/gene="At2g36040"
/notes="synonym: F11F19.5; predicted by genefinder"
<26200..26382
/gene="At2g36040"
26200..26382
/gene="At2g36040"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD21453.1"
/db_xref="GI:4510365"
/translaton="MDASDNSMKDGFNVBEFQKVEQVSKTYEKKTKPKVRSIDGDK
LGFALSSRSSESDDE"
repeat_region 26564..26660
/rpt_family="AT_rich"
repeat_region complement(28601..28628)
/rpt_family="(TA)n"
29257..29469
/gene="At2g36030"
/notes="synonym: F11F19.6; predicted by genefinder"
<29257..29469
/gene="At2g36030"
29257..29469
/gene="At2g36030"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD21454.1"
/db_xref="GI:4510366"
/translaton="WTVDEWRLLDVSTAFSGNSVPDFGESADDDAVVVLVKGBEA
VEKLSEIVFGNVLVLENHLKHGLSEI"
repeat_region 29525..29605
/rpt_family="(GGA)n"
31440..33782
/gene="At2g36020"
/notes="synonym: F11F19.7; predicted by
genefinder; supported by full length cDNA: Ceres:13310"
join(31440..31565,31654..31773,31864..31919,32009..32107,
32534..32590,32698..32761,32865..32999,33109..33192,
33299..33782)
/gene="At2g36020"
join(31886..31919,32009..32107,32534..32590,32698..32761,
32865..32999,33109..33192,33299..33604)
/gene="At2g36020"
/codon_start=1
/product="expressed protein"
/protein_id="AAD21455.2"
/db_xref="GI:20197992"

Alignment Scores:
Pred. No.: 6.63e-61 Length: 120938
Score: 999.00 Matches: 305
Percent Similarity: 42.53% Conservatave: 82
Best Local Similarity: 33.52% Mismatch: 201
Query Match: 28.04% Indels: 323
DB: 8 Gaps: 31

US-10-624-201A-2 (1-688) x AC007017 (1-120938)

Qy 22 HisGlyAsn-----SerAsnAsnAsnAsnGlnThrLeuTyLeu 35
|||||
Db 65344 CACGAAACCCCGGAGATCTCTGCCGATCCGCGGTCTTCAACAGTTGATCCTC 65403
|||||

Qy 36 MetAsnProAsnAsnTyMetGlnGlyThrThrSerAspThrGlnGlnGln 55
|||||
Db 65404 ATGAATCCAACTACTTACGTTTCAG---TACACCAACAAAGACACGACTCGAAACAAAC 65460
|||||

```

```

Qy 56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCyHisAlaAsnIle--- 74
|||||
Db 65461 -----AACACACGACCAATAGCAACAAACAAACAAACAAACAAACAAAC 65508
|||||

Qy 75 -----GlnHisAlaProLeu-----GlnGlnGln 82
|||||
Db 65509 AACACAAACAGTAGTTTCGTTTCTCGATTCCACGCGCGCGACCAACGCGAGCCAG 65568
|||||

Qy 83 HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97
|||||
Db 65569 CAGTTTCGCGAATACCACTCTCAGGTCAGAGCTGCTTCCATTACAGCGCGCGCAAC 65628
|||||

Qy 98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
|||||
Db 65629 ATCTCCGTACTTCACGGTTATCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCA 65688
|||||

Qy 117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrAsp 136
|||
Db 65689 GTG----- 65691

Qy 137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154
|||||
Db 65692 -----GATCCACTCACCAAGCGCGCG 65715

Qy 155 -----GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerProGlnLeuGlnGln 172
|||||
Db 65716 TGTGAGACGCCACGCGCGCAGCAAGGCTCTCTTTAACCTCTCTCTCAACAGCAGCAG 65775
|||||

Qy 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerSerProArgThr 187
|||
Db 65776 CAACAGCAACATCATCAACAAACACAGCCCTATTACGCTCGGATTCGGGTCGGACATGGA 65835

Qy 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202
|||||
Db 65836 GAAGATATCCGGTCCGGTCTGCTCTACAGGATCGGGGTAAACAAACGGTATAGTAAT 65895
|||||

Qy 203 MetValLeuGlySerTyLeuLeuAlaGlnGlnLeuLeuAspGluValValAsn 222
|||||
Db 65896 CTTGTT---AGTCCCAAGTACTTGAAGCAGCAGAGAGCTCTTTGACGAAGTAGTCAAC 65952
|||||

Qy 223 IleValGlyLysSerIleLysGlyAspAspGln-----LysLysAspAsnSer 238
|||||
Db 65953 GCTGATTCGATGACATGACGTAACGTAATCCCACTATTCATCGAAAAAGGGTAGTTCG 66012
|||||

Qy 239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
|||||
Db 66013 GGAATGATTAACCTGTCCGAGAAATCATCGCGCGCTCGAGGAGAGGTTCCGGTGGC 66072
|||||

Qy 258 GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
|||||
Db 66073 GGAGCAGAACGACGCCGGAACGCT-----CCGGTGAGCTAGGCACGCGCAGAGAGA 66123
|||||

Qy 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeu----- 291
|||||
Db 66124 CAAGAAATACAGATGAAGAAAGCAAACTTAGTAACATGCT-TCATGAGGTAAACAACAT 66182
|||||

Qy 291 ----- 291
Db 66183 ACGTATTATATACAATAAGCCTACATATACTAGTGTCTCTTTATATACACATGAAGTG 66242
|||||

Qy 291 ----- 291
Db 66243 GACTTAAAGTTCTGTCCCTAAGATTTTCAGTCATTTTCACAATTCATTTATATGTATA 66302
|||||

Qy 291 ----- 291
Db 66303 TTAAGCTATCCTTTTTTTTTTTTGTCTCTCCGATTTCTCAAATTGACAACTTACAAGATT 66362
|||||

Qy 291 ----- 291
Db 66363 GATGAAATCTTTTTTTTTTTTGTCCCTCTCTTAGTAAGACTAAACAACATTTTATAGATTAT 66422
|||||

```

QY	291	-----	291
Db	66423	ATTATGTAGACCGTACATCTATTGTTTCATAAGACTCACGTTACTTGTGGTTGAAAT	66482
QY	292	GluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleValLeuSer	311
Db	66483	GAAAGGTGGAGCAGAGATATAGACAGTACCACAGCAGATGCAGATGGTGATCTCTCG	66542
QY	312	PheGluGlnValAlaGlyIleGlySerAlaIleValSerTyrThrClnLeuAlaLeuHisAla	331
Db	66543	TTGCAGCAACGGCAGGGATAGGATCAGCGAAGTCATACACGTCGTAGCATGAAACCC	66602
QY	332	IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer	351
Db	66603	ATATCAAGACAGTTCGTGCTTGAAGAGCGCATCGCTGGTCAGATATAAAGCGGCCAAC	66662
QY	352	LysSerLeuGlyGluGluGlyLeuGly-----GlyLysIleGluGlySerArgLeu	369
Db	66663	AAGAGTCTTGGGAGGAAGATTCAAGTGTCTGTGTGGTGGAGGTTTGAGGGTTCGAGGCTC	66722
QY	370	LysPheValAspHisHisIleuArgGlnGlnArgAlaLeuGlnIleGlyMetMetGln	389
Db	66723	AAGTTCGTGGACCAACCATCTGAGACAGCAAGAGCTCTTCAACAACCTGGGAATGATTCAA	66782
QY	390	-----ProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerVal	406
Db	66783	CATCCTTCCAAATAATGCTTGGAGACCTCAACGTGGTCTCCGAAACGAGCCGCTCAGTT	66842
QY	407	LeuArgAlaTrpLeuPheGluHisPheLeuHisPro-----	418
Db	66843	CTCCGTGCTTGGCTCTTCGAACACTTCTTCATCCGTAAGTAACCACTATACACATG	66902
QY	418	-----	418
Db	66903	CACTACTTTTCATCACTCGGAACACAACTCAGTCTCTAAGTGTCTTCTTTTAATTAGA	66962
QY	419	TyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGln	438
Db	66963	TACCCTAAGGATTCGCAACAGCACATGCTAGCTAAGCAACAGGACTCACTCGTAGCCAG	67022
QY	438	-----	438
Db	67023	GTACGTTTCAAAATGATCTTCTTGTGTATTAGTGTACTTTAAGAAAGCGTTTAAAGT	67082
QY	439	-----ValSerAsnTrpPheIleAsnAlaArgValArgLe	450
Db	67083	AAATTTATTTCTTTAATTTATTTAGTGTGCAACTGGTTTATAACCGCAGAGTTCGGTT	67142
QY	450	uTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSe	470
Db	67143	ATGGAACCAATGTTGGAGGAGATGTACATGAGGAAATCAAGGACGAGCAAGAACAT	67202
QY	470	rThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluG1	490
Db	67203	G-----GGATCCATGGAAAGACTCTCTTGGATCAAGCAAGCAACGAAGA	67244
QY	490	uLysHisProIleThrSerSerLeuLeuGlnAspGlyIleThrThrClnAlaG1	510
Db	67245	T-----TCTGCTTC	67253
QY	510	uIleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----Hi	528
Db	67254	AAAGTCAACAAGTAAACCAAGAAAGCCCAATGGCGGACACTAATTACCATATGAATCC	67313
QY	528	sAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAs	548
Db	67314	CAATCAACCGGTGACCTAGAGGCGCTCACTGGAATGCAAGGAAGCCCAAGAGACTAAG	67373
QY	548	pHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSe	568
Db	67374	AACCAAGGAGAGACAAATGATGAGCCCAATAAT-----GCCGATTTTCACTCCAAACGA	67427
QY	568	rIleLeuSerSerValAspMetGluAlaLysAlaArgGluSerAsnLysGlyPheTh	588
Db	67428	GAAGTCTCAGTGAAAGATTCTTAGAAGACGGCAAGGATAAGATCAGATGGTGGCTAC--	67485
QY	588	rAsnProLeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp--	604
Db	67486	----CCTTTTCATGGGTAATTTTCGGCAATACCAAAATGGATGGATGTCAAGATTTTGATGT	67541
QY	605	-ProHisAspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLe	622
Db	67542	AGTCTCAGACCAAGAGCTCATGCGCAAGGTAAGTCTCAGGAAACCAACATGGGTGTCCTT	67601
QY	622	uThrLeuGlyLeuProProSerGluAsnLeuAlaMetProValSerGln-----G1	639
Db	67602	CACGTTAGTTTACCTTCATTGTGATAGCTGTGTCGTCACGCAACCATCAGGTTTTCATGCA	67661
QY	639	nAsnTyrLeuSerAsnAspLeuGlySerArgSerGluMetGly-----	653
Db	67662	GACCCACCATGGGATTCTTATAGGAGAGAGCTGAAAAATAGGAGAAACAGAGGAATATGG	67721
QY	653	-----	653
Db	67722	ACCGCCACCATCAATGTTGTTAGTTCGACCAACACCGCACATTCATCAGCGGCAGCTGC	67781
QY	654	-SerHisTyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPh	673
Db	67782	CGCGGCTTACAATGGATGAACATACAGAAC-----CAGAAGAGATA	67823
QY	673	eProThrGlnLeuLeuProAspPheVal	682
Db	67824	TGTGCTCAGTTATTTCGCCGACTTCGTT	67851
RESULT 13			
AK120844			
LOCUS			
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J023022M14, full insert sequence.		
ACCESSION	AK120844		
VERSION	AK120844.1 GI:37990467		
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	1		
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team., Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group., Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Nariikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka M., Ryu,R., Ueda M., Matsubara,K., RIKEN., Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
JOURNAL	Science 301 (5631), 376-379 (2003)		
MEDLINE	22752273		
PUBMED	12869764		
REFERENCE	2		
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,		

QY 342 AlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu-----Glu 357
 Db 1195 GCATTCGCGCGCAGGTGAGGGCGCGAGCGGGGGCTCGGGAGGACTGCGGCGAGCAGC 1254
 QY 358 GluGlyLeuGlyGly-----LysIleGluGlySerArgLeuLysPheValAspHis 374
 Db 1255 GAGGGCGCGGTGGCGGCGGAGGACGCGGTGGGTGCGGCTGCGGTTCATCGACAC 1314
 QY 375 HisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAla----- 392
 Db 1315 CAGCTCCGCGCAGCAGCGCGATGTCAGCAGCTTGGCATGTGTCAGCCCGCGCGCGGT 1374
 QY 393 -----TrpArgProGlnArgGlyLeuProGluArgAlaValSerVal 406
 Db 1375 GGGCGCGCGCGCGCGGTGGCGCGCGCAGCGTGGCTCCCGAGCGCGCGTCTCCGTC 1434
 QY 407 LeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIle 426
 Db 1435 CTCGCGCTTGGCTCTTCGAGCATTTCTCCACCCATACCCCAAGGATTCGGACAAGTC 1494
 QY 427 MetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAla 446
 Db 1495 ATGCTCGCAAGCAACCGCGCTCACAGAGCGCAGGTGCGAACTGGTTTCATCAACGG 1554
 QY 447 ArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGlnValLysAsnGln 466
 Db 1555 AGGTGAGCTGTGGAAGCAATGTGTGAGGAGATGTACGCCGAGGAGACCAAGGCCAAG 1614
 QY 467 GluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAla 486
 Db 1615 GAGGAG----- 1620
 QY 487 ProAsnGluGlyLysHisProIleThrSerSerLeuLeuGlnAspGlyIleThrThr 506
 Db 1621 ---GAGGAGGAGGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1677
 QY 507 ThrGlnAla-----GluIleSerThrSerThrIleSerThrSerPro 520
 Db 1678 GAGCAGCTCCGAGTAAGCGGAGCAGCATGTGCGCATTTGGCATGTGTCGTGCGCG 1737
 QY 521 ThrAlaGlyAlaSer-----LeuHis-----HisAlaHisAsnPhe 532
 Db 1738 GCGGTGCGCGCTCCAGGTGAGTGTGTGTCACGCGCGAGACGACGCTCAGGCGAGC 1797
 QY 533 SerPheLeuGlySerPheAsnMetAspAsn-----Thr 543
 Db 1798 TTCTACG 1857
 QY 544 ThrThrThrValAsp-----HisIleGluAsnAsnAlaLys 555
 Db 1858 ACGACACGCGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1917
 QY 556 ---LysGlnArgAsnAspMetHisLysPheSerProSerSer----- 568
 Db 1918 GTGAGCCACCGGAGCTGCTAATGAAGTTCACTGAGCGCGCGCGCGCGCGCGCG 1977
 QY 569 -----IleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLys 585
 Db 1978 GGGCACCACGCGCTCAACGACGACGACGACGCGTCCCTCGGCGCGCGCGCGGTACTCG 2037
 QY 586 GlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspPro 605
 Db 2038 CTGTTCACG-----GGCGGCGAGTACGCGGCATCAGTTCGCG-----TCGGAC 2079
 QY 606 HisAspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGly 625
 Db 2080 CATTCGCGTTCGCGCGCAGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2139
 QY 626 LeuProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAsp 645
 Db 2140 CTCGCCGACGCGCGCGCAG---CAGACGCGCGCGGTGCG-----TTCTCATAGCGCGC 2187

QY 646 LeuGlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGluAsnIleAsp 665
 Db 2188 GCGCGCGGCGAGCAGCGCGCGCGCGCGGTGACGACGCGCGGTACGAC---ATGAAC 2244
 QY 666 PheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
 Db 2245 ATGCAGAGC---ACCAAGTCGTTGGCTGCTCAGCTCATGAGAGACTTCGTG 2292
 RESULT 14
 LOCUS AK120813
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023014M03, full insert sequence.
 ACCESSION AK120813
 VERSION AK120813.1 GI:37990436
 KEYWORDS FLI_CDNA; CAP trapper.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Eriocarpaceae; Oryzaeae; Oryza.
 REFERENCE 1
 AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team.; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group.; Tsunoda, Y., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Ohtsuka, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
 TITLE Japonica rice
 JOURNAL Science 301 (5631), 376-379 (2003)
 MEDLINE 22752273
 PUBMED 12869764
 REFERENCE 2
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Nami, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Tota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.
 TITLE Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice
 JOURNAL Unpublished
 AUTHORS 3 (bases 1 to 2643)
 REFERENCE Kikuchi, S.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of

Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007). This clone is one of the 32K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M., and Nakahama, Y.

PAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsuura, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komoto, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oeato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

```

1. .2643
location/Qualifiers
  organism="Oryza sativa (japonica cultivar-group)"
  mol_type="mRNA"
  cultivar="Nipponbare"
  db_xref="taxon:39947"
  clone="J023014M03"

```

ORIGIN

Alignment Scores:	
Pred. No.:	3.34e-59
Score:	946.00
Percent Similarity:	46.03%
Best Local Similarity:	36.28%
Query Match:	26.55%
DB:	8
Gaps:	34
Indels:	206
Mismatches:	215
Conservative:	76
Matches:	283
Length:	2643

```

US-10-624-201A-2 (1-688) x AK120813 (1-2643)
Qy      2 TyrTyrGlnGlyThrSerAspAsnThrAsn-----IleGlnAlaSpHisGlnGlnArgHis 20
|||::: |||::: |||::: ::|||::|
Db     343 TACTTCACAGCGCGCGCGCACGACGTCTGCAGGCC----- 384
|||||::| |::|::| |::|::| |::|::|
Qy      21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
|||::: |||::: |||::| |::|::| |::|::| |::|::|
Db     385 -----GCCGCGCACCAGCGGCTGCAGAGCGCTTACCTCATGAACCCC---AGC 429
|||::: |||::: |||::| |::|::| |::|::| |::|::|
Qy      41 TyrMetGlnGlyTyrThrThrSerAspThr-----GlnGlnGlnGlnGln 55
|||::: |||::: |||::| |::|::| |::|::|
Db     430 TACGTC---GGCTTCACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCCCAAC 486
|||::: |||::: |||::| |::|::| |::|::|
Qy      56 LeuLeuPheLeuAsnSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGln 75
:::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     487 ATGGTGTTCTCAACTCG-----GCCGTGAGCACCCTTCACGCGCGGAGGCTTCAGC 537
|||::: |||::: |||::| |::|::| |::|::| |::|::|
Qy      76 His-----AlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAla 91
|||::: |||::: |||::| |::|::| |::|::| |::|::|
Db     538 CACCACCACCACCGCACACCG---GCCGCGCAGCAGCACTTCGTCGCATCCCTCTCCAGTCG 594
|||::: |||::: |||::| |::|::| |::|::| |::|::|
Qy      92 ValSerLeuHisAspGlnIleAsnHisGlyLeuLeuGlnArgMetTrpAsnAsnGln 111

```

Db	595	GGG-----	::::	-TACAACCTGTGGGCG-	612
Qy	112	AspGlnSerGlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGly	 ::: :::		131
Db	613	-----	CCTGAGCCACCAGCGGCAGCATGTGCG-		642
Qy	132	GlyIleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGln	 ::: :::		151
Db	643	-----	CGCCCAACGCG		651
Qy	152	HisArgGlnGlnGlnGln-----	GlnGlnGlyGlyLeuSerLeuSerLeuSerLeuSerPro	168	
Db	652	CAGCGGCCCCAGCAGCAAGCCAGCAGCAGTCGGCACATGTCGGCGCGCGGTGACGCCCC	711		
Qy	169	GlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsn	188		
Db	712	GTGCTGAGC-----	CTGTCGTGCGGTGAGCGCGCGCG	744	
Qy	189	AsnValThrIleArgGlyThrLeuAspGlySerSer-----	Ser	201	
Db	745	CCGCTCACCTGGCCCGCGCGCGCGCGCGCTCCC	CGCGCGCACCGACACGAG	804	
Qy	202	AsnMetValLeuGlySerLysTyrluLysAlaAlaGlnIleuLeuAspGluValVal	221		
Db	805	AAGGTGGTGATCGCTCCAGGTACTCGAAGCGCGCGCAGGAGTGTCTCGACGAGCGGTG	864		
Qy	222	AsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsnLys	241		
Db	865	AGC---GTCAGCAAGGGCGCGCGCGCGCGGTGAAGNAGNAGGAGNACTCGAGGGCGCG	921		
Qy	242	GluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSer	261		
Db	922	GTGTCC-----	GGTGGCGCGCGGTGCGGAGGACGGT	954	
Qy	262	SerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGln	281		
Db	955	GGCGCGCAGCAAGAGCGCGCGCGCGCGAGATCGACGCGCGAGCGCGCAGSAGCTGCAG	1014		
Qy	282	MetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrluArgGlnTyrlu	301		
Db	1015	ATGAGNAGAGCAAGTACTGTAACATGCTCGATGAGGTGGAGCAGCGGTACACGCGAGTAC	1074		
Qy	302	HisHieGlnMetGlnIlelleValLeuSerPheGluGlnValAlaGlylleGlySerAla	321		
Db	1075	CACCGCGAGATGACGGGGTGGCGCGCGCGTTCGACGCGCGCGCGCGCGGTGCGCG	1134		
Qy	322	LysSerTyrluThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAsp	341		
Db	1135	ACACAGCTACAGTCGTGGCGCTGCGCACCATCTCGCGCAGTTCCGGTGCCTCCCGCAC	1194		
Qy	342	AlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu-----	Glu	357	
Db	1195	GCCATCGCGCGCAGGTGACGGCGCGCAGCCCGGGGCTCGGGAGGAGNACTCGCGCGAC	1254		
Qy	358	GluGlyLeuGlyGly-----	LysIleGluGlySerArgLeuLysPheValAspHis	374	
Db	1255	GAGGGCGCGGTGCGCGCGGAGACACACGCGTGGGTGCGCGTTCATCGACCCAC	1314		
Qy	375	HisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAlaAla-----	392		
Db	1315	CAGTCCGGCAGCAGCGCGCATGTCAGCAGCTTGGCATGTGTGACGCGCGCGCGCGGT	1374		
Qy	393	-----	TTrpArgProGlnArgGlyLeuProGluArgAlaValSerVal	406	
Db	1375	GGCGCGCGCGCGCGGTGGCGCGCGCAGCGTGGGTCTCCCGAGCGCGCGCTCGTC	1434		
Qy	407	LeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrluProLysAspSerAspLysIle	426		
Db	1435	CTCGCGCTGGCTCTTCGAGCATTTCTCCACCCATACCCCAAGGATTTCGACAGGTC	1494		
Qy	427	MetLeuAlaLysGlnThrGlyLeuThrArgSerGln-----	438		

Db 1495 ATGCTCGCAAGCAACCGCCTCACCAGGAGCGAGTACGTACGTATTACGTCTACGTAC 1554
QY 439 -----ValSerAsnTrpPhe 443
Db 1555 AATACTACGTCTACGTATGTTTGAAGTACGTGTACGTATTGTTAGGTGTGCAACTGGTTC 1614
QY 444 IleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluVal 463
Db 1615 ATCAAGCGAGGGTGGGTGTGGAGCAATGTTGGAGAGATGTATGCCGAGGAGACC 1674
QY 464 LysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsn 483
Db 1675 AAGGCCAAGGAGAG----- 1689
QY 484 IleSerAlaProAsnGluLysHisProIleIleThrSerSerLeuLeuGlnAspGly 503
Db 1690 -----GAGGAGAGGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1737
QY 504 IleThrThrThrGlnAla-----GluIleSerThrSerThrIleSer 517
Db 1738 GCGTGGCGGAGCAGCTCGAGTAGCCGAGCAGAGTGGCGATTCGCGTTCGTCG 1797
QY 518 ThrSerProThralaGlyAlaSer-----LeuHis-----HisAla 529
Db 1798 TCGTCGCCGCGTGGCGGCTCAGGTGAGTGTGTCAGCGCCGCGGAGCAGCAGCGCT 1857
QY 530 HisAsnPheSerPheLeuGlySerPheAsnMetAspAsn----- 542
Db 1858 CAGGCGAGCTTTCACG 1917
QY 543 -----ThrThrThrValAsp-----HisIleGluAsn 552
Db 1918 AAGCGCAGCAGCAGCAGCGCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1977
QY 553 AsnAlaLys-----LysGlnArgAsnAspMetHisLysPheSerProSerSer 568
Db 1978 GAGCGCGCGCTGAGCCAGCGGAGCTGCTAATGAAGTTCACTGAGCGCGCGCGCGCG 2037
QY 569 -----IleLeuSerSerValAspMetGluAlaLysAlaArgGluSer 582
Db 2038 GTGAGAACGGCGCCACCGCAGCTCAACGAGCAGCAGCAGCAGCTCCTCGCGCGCGCGCG 2097
QY 583 SerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArg 602
Db 2098 GGGTACTCGCTGTTACG-----GCGCGCGCAGTACGGGCATCAGTTCGCGC--- 2142
QY 603 PheAspProHisAspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeu 622
Db 2143 ---TCGACCATTTTCGCTTCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCTC 2199
QY 623 ThrLeuGlyLeuProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeu 642
Db 2200 AGCTCGGCTTCCCGCAGCGCGCGCGCAGC---CAGACGCGCGCGCGCGCGCGCGCGCG 2247
QY 643 SerAsnAspLeuGlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGlu 662
Db 2248 ATAGCTC 2307
QY 663 AsnIleAspPheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
Db 2308 ---ATGAACATGCAGAGC---ACCAAGTCGTTGGCTCGCTCAGCTCATGAGACTTCGTG 2361
RESULT 15
AK070465
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023054P21, full insert sequence.
ACCESSION AK070465
VERSION AK070465.1 GI:32980489
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nani, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otono, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayaehizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)
2752273
12869764
2 (bases 1 to 3030)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayaehizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, T., Nakamura, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nani, K., Niikura, J., Nani, K., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otono, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression: 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007]
This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cdna/NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nani, K., Ohtsuka, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Maeda, H., Miura, J., Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayaehizaki, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,

33	LeuTyrLeuMetAenProAsnAenTyrMetGlnGlyTyrThrThr-----SerAspThrGln	51
897	CTCTATCTGAACAATCCATCTTCT-----GGACCTTATACGGAGTTTCAGCGGCATTTCTG	950
52	GlnGlnGlnGlnLeuPheLeuAenSerSerProAlaAlaSerAenAlaLeuCysHis	71
951	CAGCTCAGCAGAAATTCATGGAGATCCCTGGCCCTGGG-----	989
72	AlaAsnIleGlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAla	91
990	-----CATGCGTCAGCAATGTCACAAGAC-----CCCTCT	1019
92	ValSerLeuHisAspGlnIleAenHisHisGlyLeuLeuGlnArgMetTrpAenAsnGln	111
1020	TCAAGGAGATCCGACATCTTCTTCGCACACGGGTCAACGCTTTTCAGCCCATGTCAAA	1079
112	AspGlnSerGlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGly	131
1080	GATATGAAATGAGATGTTGATGCATATGATGATGAGACACAAAGTAGTGGCTCTGAG	1139
132	GlyIleThrThrAsp-----LeuAlaSerGlnLeuAlaPhe-----	143
1140	CTCATCCATGATGATGCCACACATGGTTTCACAGCTTGAGTTTGGTGTCTGAATAATCAC	1199
144	---GlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlyLeu	162
1200	AACTCATCAAGTGTTCATCA-----ATGCAAAGCCCAAGGGCTA	1238
163	SerLeuSerLeuSer-----Pro	168
1239	TCTCTGAGCTTGAACACACAAATCATCGCGCTTCCTTACGTTACTGGTCTATAAAGCCA	1298
169	GlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsn	188
1299	GACATGTTAACGGCGGAGTCTTACCATGATCAACCTTAGAGGGGAAGACATGCGGATGAAG	1358
189	AsnValThrIleArgGlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLys	208
1359	AAT-----TTGCAGTCGGAAGCTTCACGCGCAATCCCGAACTCAAGG	1400
209	TyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsnIleValGlyLysSerIle	228
1401	TATCTCAAGCAGCACAAAGAACTGCTTGATGAGTCTGTGAGT---GTTTGGAGAGACATA	1457
229	LysGlyAspAspGlnLys-----LysAspAsnSerMetAsnLysGlu	242

```

QY      595 TyrAlaMetGlyAspPheGlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPhe 614
Db      2484 TATCACTTGGCAGAGCTTGGAGATAC----- 2510
QY      615 HisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProProSerGluAsnLeuAlaMet 634
Db      2511 ---GGCAATGGCAATGTGCTGACACTGGGCTTCAGCACTCCAGCAGCAACCTCGTT 2567
QY      635 ProValSerGlnGlnAsnTyr-----LeuSerAsnAsp-----LeuGlySerArgSer 650
Db      2568 CCTAATGCTCAGCCAGGTTTCCCTGGTGTAAATGAGGATGACATCTACAATGCTACTGCT 2627
QY      651 GluMetGlySerHisTyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsn 670
Db      2628 CCTCTTGGGTGTCACCGTCGCATCTTCGACTATGATTCGATGAATCAGATGGATCAACGG 2687
QY      671 LysArgPhePro---ThrGlnLeuLeuProAspPheVal 682
Db      2688 CAACGGTTTGAGCATTACCTCTTCTGCAATGATTTTGTG 2726

```

Search completed: July 29, 2005, 17:21:49
Job time : 8239 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 08:33:52 ; Search time 949 Seconds
(without alignments)
4291.657 Million cell updates/sec

Title: US-10-624-201A-2

Perfect score: 3563

Sequence: 1 MYQGSTDNITQADHQRRH.....GNKRFFPTQLLPDFVTGNLGT 688

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10624201/runat_27072005_122956_10897/app.query.fasta_1.839
-DB=N_Geneseq_16Dec04 -QFMR=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US10624201.CGN 1.1.644 @runat_27072005_122956_10897 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_16Dec04:*
1: Geneseq_1980s:*
2: Geneseq_1990s:*
3: Geneseq_2000s:*
4: Geneseq_2001as:*
5: Geneseq_2001bs:*
6: Geneseq_2002as:*
7: Geneseq_2002bs:*
8: Geneseq_2003as:*
9: Geneseq_2003bs:*
10: Geneseq_2003cs:*
11: Geneseq_2003ds:*
12: Geneseq_2004as:*
13: Geneseq_2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1150.5	32.3	2043	6 ABZ13532	Abz13532 Arabidops
2	1150.5	32.3	2043	9 ADB23149	Adb23149 Environme
3	1148.5	32.2	2385	10 AD030303	Add30303 Plant yle
4	1148.5	32.2	2385	12 ADI44208	Adi44208 Plant tra
5	926.5	26.0	1929	12 ADQ37104	Adq37104 Cell prol

6	826	23.2	1971	12	ADO62706	Ado62706 Transcrip
7	824.5	23.1	1617	12	ADO63042	Ado63042 Transcrip
8	824	23.1	1884	6	ABZ13100	Abz13100 Arabidops
9	824	23.1	1983	4	AAD06494	Aad06494 Arabidops
10	824	23.1	1983	10	ADc46618	Adc46618 Thalecres
11	824	23.1	1983	10	ADD30786	Add30786 Plant yle
12	824	23.1	1983	12	ADI44018	Adi44018 Plant tra
13	824	23.1	1983	12	ADO02224	Ado02224 Thalecres
14	820	23.0	2025	3	AAC39138	Aac39138 Arabidops
15	816	22.9	2458	12	ADO62707	Ado62707 Transcrip
16	803	22.5	1824	3	AAC36745	Aac36745 Arabidops
17	802	22.5	1449	3	AAC43406	Aac43406 Arabidops
18	802	22.5	1449	6	ADG88236	Adg88236 A. thalia
19	802	22.5	1575	4	AAD06493	Aad06493 Arabidops
20	802	22.5	1575	12	ADO61968	Ado61968 Transcrip
21	802	22.5	1599	6	ABZ14349	Abz14349 Arabidops
22	802	22.5	2031	3	AAC36460	Aac36460 Arabidops
23	797.5	22.4	2545	4	AAD06479	Aad06479 Arabidops
24	797.5	22.4	2545	10	ADE37296	Ade37296 Plant yle
25	797.5	22.4	2545	12	ADI43866	Adi43866 Plant tra
26	797.5	22.4	2545	12	ADO01854	Ado01854 Thalecres
27	786	22.1	1485	3	AAC50215	Aac50215 Arabidops
28	732.5	20.6	1836	6	ABZ12846	Abz12846 Arabidops
29	732.5	20.6	1836	6	ADG88038	Adg88038 A. thalia
30	725	20.3	2405	4	AAD06449	Aad06449 Arabidops
31	725	20.3	2405	10	ADE37116	Ade37116 Plant yle
32	725	20.3	2405	12	ADI43868	Adi43868 Plant tra
33	725	20.3	2405	12	ADO01856	Ado01856 Thalecres
34	701.5	19.7	1296	12	ADO61794	Ado61794 Transcrip
35	678	19.0	2034	3	AAC50226	Aac50226 Arabidops
36	649.5	18.2	1905	3	AAC39451	Aac39451 Arabidops
37	646.5	18.1	2442	10	AAL55526	Aal55526 qSH-1 gen
38	646.5	18.1	2450	10	AAL55527	Aal55527 qSH-1 gen
39	639	17.9	894	12	ADQ37116	Adq37116 Cell prol
40	636	17.9	706	6	ABK82114	Abk82114 DNA encod
41	616.5	17.3	2131	6	ABL56175	Abl56175 Hordeum v
42	607	17.0	2141	6	ABL56177	Abl56177 Hordeum v
43	605	17.0	626	5	AAH87740	Aah87740 Peppermin
44	586.5	16.5	1239	3	AAC55934	Aac55934 Eucalyptu
45	493	13.8	529	3	AAC55946	Aac55946 Eucalyptu

ALIGNMENTS

RESULT 1	
ABZ13532	
ID	ABZ13532 standard; DNA; 2043 BP.
XX	
AC	ABZ13532;
XX	
DT	21-JAN-2003 (first entry)
XX	
DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 1337.
XX	
KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO200216655-A2.
XX	
PD	28-FEB-2002.
XX	
PF	24-AUG-2001; 2001WO-US026685.
XX	
PR	24-AUG-2000; 2000US-0227866P.
PR	26-JAN-2001; 2001US-0264847P.
PR	22-JUN-2001; 2001US-0300111P.
XX	
PA	(SCRI) SCRIPPS RES INST.
PA	(SVGN) SYNGENTA PARTICIPATIONS AG.
PI	Harper JF, Kreps J, Wang X, Zhu T;
XX	

DR WPI; 2002-304127/34.
XX Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
XX
PS Claim 144; SEQ ID NO 1337; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB21196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
XX
SQ Sequence 2043 BP; 627 A; 507 C; 507 G; 402 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1-32e-85 Length: 2043
Score: 1150.50 Matches: 306
Percent Similarity: 51.81% Conservative: 81
Best Local Similarity: 40.96% Mismatches: 201
Query Match: 32.29% Indels: 159
DB: 6 Gaps: 28

US-10-624-201A-2 (1-688) x AB213532 (1-2043)
QY 22 HisGlyAsn-----SerAsnAsnAsnIleGlnThrLeuTyLeu 35
DB 16 CACGGAAACCCACCGAGATCTCTCGCGATCCGACGGTGTCTTCAAAAGTTGATCCTC 75
QY 36 MetAsnProAsnAsnTyMetGlnGlyTyThrThrSerAspThrGlnGlnGlnGln 55
DB 76 ATGAATCCAACTACTTACGTTCTAG---TACACCCAAACAGACAAACGACTCGACACAAAC 132
QY 56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaIle--- 74
DB 133 -----AACAAACGACAAATAGCAACAAACAAACAAACAAACAAACAAACAAAC 180
QY 75 -----GlnHisAlaProLeu-----GlnGlnGln 82
DB 181 AACAAACAGTAGTTTCTGATTCCTCGATTCCTCCGCGCGCGCCGACCAACGCGAGCCAG 240
QY 83 HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97
DB 241 CAGTTTCGCGAATACCACTCTCAGGTCACGAGCTGCTTCCATTACAGCCGCGGACAAAC 300
QY 98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
DB 301 ATCTCGGTACTTCACGGTTATCTCCGCGCGGTGCTAGTACAGTCTCTACGCGTAGCCACCAA 360
QY 117 ValIleValProSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
DB 361 GTG----- 363
QY 137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154
DB 364 -----GATCCCACTCACCGACGACGCGCG 387
QY 155 -----GlnGlnGlnGlnGlnGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
DB 388 TGTGAGACGCGACGCGCGAGCAAGCGCTCTCTTAACCTCTCTGCTCTCAACAGCAGCAG 447
QY 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerProArgThr 187
DB 448 CAACAGCAACATCATCAACACACACGACCTATTTCAGTCGTGGATTCGGGTCCGAGACATGGA 507
QY 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202

508 GAAGATATCCGGTCCGGTCTCGCTCTACAGGATCGGGGTAAACAAACGGTATAGTAAAT 567
QY 203 MetValLeuGlySerIleTyLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
DB 568 CTTGTT---AGCTCCAAAGTACTTGAAGGCGACACAAAGAGCTTCTTGACGAAGTAGTCAAC 624
QY 223 IleValGlyLysSerIleLysGlyAspAspGln-----LysLysAspAsnSer 238
DB 625 GCTGATTCGATGACATGAACCTAAATCCCAACTATTCTTCGAAAAGGGTAGTTGC 684
QY 239 MetAsnGlyLysSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
DB 685 GGAATATGATAACCTCTCGGAAATCATCGCGCGCGTGGAGGAGAGTTCCTCGGTGC 744
QY 258 GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
DB 745 GGAGCAGAGACGACCGCGGAAACGT-----CCGGTGGAGCTAGGCACGCGCAGAGAGA 795
QY 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGlnArg 297
DB 796 CAAGAAATACAGATGAAGAAAGCAAACTTAGTAACATGCTTTCATGAGTGGAGCAGAGA 855
QY 298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
DB 856 TATAGACAGTACCACAGCAGATGCAGATGGTGTATCTCTTCGTCGAGCAAGCGCAGGG 915
QY 318 IleGlySerAlaLysSerTyThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
DB 916 ATAGGATCAGCGAAGTCTATACACGTCGCTAGCATTTGAAAACCATATCAACAGACAGTTCGT 975
QY 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu 357
DB 976 TCTTGAAGAGCGCATCTGTCGATGAATAAAGCGCCCAACAGAGTCTTGGGGAGGAA 1035
QY 358 GlyGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
DB 1036 GATTTCAGTCTGTTGTTGGAGGTTTGAAGGCTCGAGGCTCAAGTTCGTGACCAACCAC 1095
QY 376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla 392
DB 1096 TTGAGACACAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCTTCCAAATATGCT 1155
QY 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
DB 1156 TGAGACCTCAAGTGTCTCCAGACAGCGGTCTCAGTTCCTCGTCTGCTTCTTC 1215
QY 413 GluHisPheLeuHisProTyProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
DB 1216 GAACACTTTCTTTCATCCATACCTAAGGATTTCGACACAGCAGATGCTAGTAAAGCAACA 1275
QY 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
DB 1276 GGACTCAGTCGTAGCCAGGTGTGAACTGGTTATATAACCGCAGAGATTCGGTTATGAAA 1335
QY 453 ProMetValGluGluMetTyLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
DB 1336 CCAATGGTGGAGAGATGTACATGGAGGAAATGAAGGACGAGCAGCAAGAACATG----- 1389
QY 473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluLysHis 492
DB 1390 -----GGATCCATGGAAGAGACTCTTTGGATCAAAAGCAACGAAGAT----- 1431
QY 493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSer 512
DB 1432 -----TCTGCTTCAAGTCA 1446
QY 513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis 530
DB 1447 ACNAGTAAACCAAGAAAGACCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1506
QY 531 AsnPhePheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550


```
QY 258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db 745 GGAGCAGAGCAGCCGGGAACGT-----CCGGTGAGCTAGGCAGCGCAGAGAGA 795
QY 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluValGluGlnArg 297
Db 796 CAAAGAAATACAGATGAAGAAGCAAAACTTAGTAACATGCTTCATCAGGTGGAGCAGAGA 855
QY 298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
Db 856 TATAGACAGTACCACCAGCAGATGCAGATGCTCTCTTCGTTCCAGCAAGCGCAGGG 915
QY 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db 916 ATAGGATCAGCGAAGTCATACAGTCGCTAGCAATGAAAACCATATCAAGACAGTTCCT 975
QY 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 976 TGCTTGAAAGAGCGGATCGCTGTCAGATAAAGCGGCCCAAGAGTCTTGGGGAGGAA 1035
QY 358 GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
Db 1036 GATTCACTGCTGCTGTGGAGGTTTGGGGGTCGAGGCTCAAGTTCGTGGACCACCCAC 1095
QY 376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla 392
Db 1096 TTGAGACAGCAAGAGCTCTTCAACACTCGGGAATGATCAACATCCTTCCAAATATGCT 1155
QY 393 TtpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db 1156 TGGAGACCTCAAGTGGTCTCCAGAACGAGCGCTCTCAGTTCCTCGTGTGGCTCTTC 1215
QY 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 1216 GAACACTTTCTTCATCCATACCTTAAGGATTCGACAAGCACAATGCTAGTAAACA 1275
QY 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db 1276 GGACTCACTCGTAGCCAGGTGTCGAACCTGTTTATAAACGAGAGTTTCGTTATGAAA 1335
QY 453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
Db 1336 CCAATGGTGGAGAGATGTACATGGAGGAAATCAAGGAGCAGGCAAGAACATG----- 1389
QY 473 ThrSerGlyAspAsnLysLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
Db 1390 -----GGATCCATCGAAGAGACTCTTTGGATCAAAAGCAACGAAGAT----- 1431
QY 493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSer 512
Db 1432 -----TCTGCTTCAAGTCA 1446
QY 513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis 530
Db 1447 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTTACCATATGAATCCCAATCAC 1506
QY 531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
Db 1507 AACGGTGACTAGAACGGCTCACTGGAATGCAAGGAAGCCCCAAGAGACTAAGAACCCAGC 1566
QY 551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
Db 1567 GACGACACATGATGACGCAATAAT-----GCGGATTTCACTCCACGAGAGACTC 1620
QY 571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db 1621 ACGATGAAGATTCTAGAAGAACCGCAAGGAGGATAAGATCAGATGCTGGCTAC-----CCT 1674
QY 591 LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db 1675 TTCATGGGTAATTTCGGGCAATACCAATGGATGAGATGCAAGATTTGATGTAGTCTCA 1734
```

```
QY 607 AspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLeuThrLeu 624
Db 1735 GACGAGAGCTCATGGCGCAAAAGTACTCAGGAAACAACAATGGCGTGTCCCTCAGTTA 1794
QY 625 GlyLeuProSerSerGluAsnLeuAlaMetProValSerGln-----GlnAsnTyr 641
Db 1795 GGTTTACCTCACTTGTGATAGCTTGTCTCCAGCACCACATCAGGGTTTCATGCAAGCCAC 1854
QY 642 LeuSerAsnAspLeuGlySerArgSerGluMetGly----- 653
Db 1855 CATGGGATTCCTATAGGGAGAAAGAGTGAATAATAGGAGAAACAGAGGAATATGGACCCGCC 1914
QY 654 -----SerHis 655
Db 1915 ACCATCAATGGTGGTAGCTCGACCACCAACCGCACATTCATCAGCGCAGCTGCCGGGCT 1974
QY 656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675
Db 1975 TACAATGGGATGAACATACAGAAC-----CAGAAAGAGATATGTGGCT 2016
QY 676 GlnLeuLeuProAspPheVal 682
Db 2017 CAGTTATTGCCCGACTTCGTT 2037
RESULT 3
ADD30303
ID ADD30303 standard; cDNA; 2385 BP.
XX
AC ADD30303;
XX
DT 15-JAN-2004 (first entry)
XX
DE Plant yield-related polynucleotide clone G1589.
XX
KW ds; transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.
XX
OS Arabidopsis thaliana.
XX
PN WO2003013227-A2.
XX
PD 20-FEB-2003.
XX
PF 09-AUG-2002; 2002WO-US025805.
XX
PR 09-AUG-2001; 2001US-0310847P.
PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-038692P.
PR 14-JUN-2002; 2002US-00171468.
XX
(WEND-) MENDEL BIOTECHNOLOGY INC.
XX
Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE,
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
PI Broun PE;
XX
WPI: 2003-248221/24.
DR P-PSDB; ADD30304.
XX
New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
Disclosure; SEQ ID NO 332; 454pp; English.
XX
The invention relates to a number of isolated Arabidopsis thaliana cDNA
sequences and their encoded proteins which are especially transcription
factor related cDNA's and proteins. The isolated or recombinant plant
transcription factor polynucleotides and polypeptides are useful in
producing transgenic plants with commercially valuable properties, i.e.
modified or altered desirable traits as compared to a reference plant,
XX such as an alteration in a plant growth characteristic, e.g. growth rate,
```

CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the cDNAs of the invention.

XX
SQ Sequence 2385 BP; 748 A; 578 C; 565 G; 494 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,39e-85 Length: 2385
Score: 1148.50 Matches: 306
Percent Similarity: 51.67% Conservatism: 80
Best Local Similarity: 40.96% Mismatches: 202
Query Match: 32.23% Indels: 159
DB: 10 Gaps: 28

US-10-624-201A-2 (1-688) x ADD30303 (1-2385)

QY 22 HisGlyAsn-----SerAsnAsnAsnAsnGlnThrLeuTyrLeu 35
DB 194 CACGAAACCCACCGGAGATCTCTGCCGGATCCGAGCGTGTCTTCAACGTTGATCCTC 253
QY 36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln 55
DB 254 ATGAATCCAACTACTTACGTTTCAG---TACACCAACAAGACCAACGACTCGAACAAC 310
QY 56 LeuLeuPheLeuAsnSerProAlaAlaSerAsnAlaLeuCyHisAlaAsnIle--- 74
DB 311 -----AACACGCAACAATAGCAACACACACACACACACACACACACACAC 358
QY 75 -----GlnHisAlaProLeu-----GlnGlnGln 82
DB 359 AACACACACAGTAGTTTCCTGATTCCTCCAGCGCGCGAGCCAAACGCGAGCCAG 418
QY 83 HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97
DB 419 CAGTTCTCGGAATACCACTCTCAGGTACGAGAGCTGTCTCCATTTACAGCCGCGACAA 478
QY 98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
DB 479 ATCTCGTACTTACCGTTATCTCCGCGGTGCAGTACAGTCTCTACGTTAGCCACCA 538
QY 117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrAsp 136
DB 539 GTG----- 541
QY 137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154
DB 542 -----GATCCCACTCACCGAGCAAGCGCG 565
QY 155 -----GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerProGlnLeuGlnGln 172
DB 566 TGTGAGACCGCAGCGCGGAGAGGCTCTCTTTAACCTCTCTGCTCTACAGAGAGCAG 625
QY 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerSerProArgThr 187
DB 626 CAACAGCAACATCATCAACACACACAGCCCTATTCACTCGGATTCGGGTCCGACATGGA 685
QY 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202
DB 686 GAAGATATCCGGTCCGGTCTGCTCTACAGGATCGGGGTAAACAACCGGTATAGCTAAT 745
QY 203 MetValLeuGlySerIleValLeuAlaGlnGlnLeuLeuAspGluValValAsn 222
DB 746 CTTGTT---AGCTCCAAGTACTTGAAGGACGACACAGAGCTTCTTGAAGAGTAGTCAAC 802
QY 223 IleValGlySerIleValLeuGlyAspGln-----LysIleAspAsnSer 238
DB 803 GCTGATTCGATGACATGACGCTAAATCCCACTATTCTCATGAAAAAGGGTAGTTCG 862
QY 239 MetAsnIleGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
DB 863 GGAATGATAAACCTGTCCGAGAAATCATCCGCGCGCGCTGGAGGAGAGGTTCCCGTGGC 922

QY 258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
DB 923 GGAGCAGAAGCAGCCGCGAAACGT-----CCGGTGGAGCTAGGCACGCGCAGAGAGA 973
QY 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
DB 974 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 1033
QY 298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
DB 1034 TATAGACAGTACCACCGAGATGCGAGATGGTGATCTCTCTGTTTCAGCAAGCGCGAGGG 1093
QY 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
DB 1094 ATAGGATCAGCGAAGTCATACAGCTCGCTAGCAATTGAAAACCATATCAACAGACAGTTC 1153
QY 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
DB 1154 TGGTTGAAAGAGCGGATCGCTGCTCAGATAAAGACGCGCCCAAGAGATCTTTGGGAGGAA 1213
QY 358 GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
DB 1214 GATTCAGTGTCTGGTGTGGAGGTTTGAGGGGTGAGGCTCAAGTTCGTGGACCCAC 1273
QY 376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla 392
DB 1274 TTGAGACACAAAGAGCTCTTCAACAACATGGGAATGATTCACACATCTCTCCAATATGCT 1333
QY 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
DB 1334 TGGAGACCTCAACGTGTCTCCAGAACGAGCGCTCTCAGTTCCTCGTGTGCTGCTTC 1393
QY 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
DB 1394 GAACTCTTCTTCATCCATCCCTAAGGATTCGACAGCACAATGCTAGTAGTAAGCAACA 1453
QY 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
DB 1454 GGACTCACTCGTAGCCAGGTGTGCAACTGGTTTATAACGCGAGAGTTCGGTTATGAAA 1513
QY 453 ProMetValGluGluMetTyrLeuGluValLysAsnGlnGlnGlnAsnSerThrAsn 472
DB 1514 CCAATGCTGGAGAGATGTACATGGAGGAATGAAGGAGCAGCAAGAACATG----- 1567
QY 473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluLysHis 492
DB 1568 -----GGATCCATGGAAAGAGCTCTTTGGATCAAAGCAACGAAGAT----- 1609
QY 493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSer 512
DB 1610 -----TCTGCTTCAAAAGTCA 1624
QY 513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis 530
DB 1625 ACAAGTAACCAAGAAAGAGCCCAATGGCGGACACTAATATACCATATGATCCCAATAC 1684
QY 531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle 550
DB 1685 AACGGTGCTAGAGCGCTCACTGGAATGCAAGGATGCCCCAAGAGACTTAAGAACCCAGC 1744
QY 551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerIleLeu 570
DB 1745 GACGAGACAATGATGACGCCCAATAAT-----GCGGATTTCACTCCCAACGAGAGCTC 1798
QY 571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
DB 1799 ACCATGAAGATTCTAGAGAACCGCAAGGGATAAGATCAGATGGTGGCTAC-----CCT 1852
QY 591 LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
DB 1853 TTCATGGGTAAATTTCCGGGCAATACCAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1912


```
QY 155 -----GlnGlnGlnGlnGlnGlnGlyLeuSerLeuSerProGlnLeuGlnGln 172
Db 566 TGTGAGAGCGCCAGCGCGCAGCAAGGCTCTCTTTAACCTCTCGTCTCAACAGCAGCAG 625
QY 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerSerProArgThr 187
Db 626 CAAAGCAACATCATCAACCAACACCGCCATTCACTCGGATTCGGGTCCGACATGGA 685
QY 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202
Db 686 GAAGATATCCGGTCCGGTCTGCTCTACAGGATCGGGGTAAACAACGCTATAGCTAAT 745
QY 203 MetValLeuGlySerIleValLeuLeuLeuAlaGlnGlnLeuLeuAspGluValValAsn 222
Db 746 CTTGTT---AGCTCCAAGTACTTGAAGGACGACCAAGAGCTTCTTGACGAAGTAGTCAAC 802
QY 223 IleValGlyLysSerIleValGlyAspAspGln-----LysLysAspAsnSer 238
Db 803 GCTGATTCGATGACATGAACGCTAAATCCCACTATTCTCATCGCAAAAGGCTAGTGC 862
QY 239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
Db 863 GGAATGATAAACCCTGTCGAGATCATCGCGCGCTCGAGGAGAGGTTCCGGTGGC 922
QY 258 GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db 923 GGAGCAGAAGCAGCGCGGAACGCT-----CCGGTGAGCTAGGACCGCAGCAGAGA 973
QY 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
Db 974 CAAGAATAACAGATGAAGAAGCAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 1033
QY 298 TyrArgGlnThrHisGlnMetGlnIleValLeuSerPheGluGlnValAlaGly 317
Db 1034 TATGACAGTACCAACGACAGATGCAGATGGTGATCTCTTCGTCGACGAGCGGAGGG 1093
QY 318 IleGlySerAlaLysSerThrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db 1094 ATAGGATCAGCAAGTCATACAGCTCGCTAGCATTTGAAACCATATATCAAGACAGTTCG 1153
QY 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 1154 TGCTTGAAGAGCGATCGCTGTGTGATGAAGACCGCCCAACAGAGTCTTTGGGGAGAA 1213
QY 358 GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
Db 1214 GATTCAGTGTCTGGTGTGGAGGTTGAGGGTTCGAGGCTCAAGTTCGTGGACCCAC 1273
QY 376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla 392
Db 1274 TTGAGACAGCAAGAGCTCTTCAACAACTGGGAATGATTCAACATCTCTTCCAATAATGCT 1333
QY 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db 1334 TGAGACCTCAACGTGCTCCCGAGAACGAGCGCTCTCAGTTCCTCGCTGCTGCTTC 1393
QY 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 1394 GAACACTTCTTATCATACCTTACGATTTCGACACGACACATGCTAGCTAGCAACA 1453
QY 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db 1454 GGACTCACTCGTACCGAGGTGTGCACTGTTTATAAACCGAGAGTTCGTTTATGAAA 1513
QY 453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
Db 1514 CCAATGCTGGAGGAGATGTATCATGTGAGGAAATGAAGGAGCAGCAAGAACATG----- 1567
QY 473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
Db 1568 -----GGATCCATGGAAGAGACTCTTTGGATCAAGACAGCAAGAT----- 1609
```

```
QY 493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512
Db 1610 -----TCTGCTTCAAAGTCA 1624
QY 513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis 530
Db 1625 ACAAGTAACAAGAAAGAGCCCAATGGCGGACACTAATTTACCATATGAATCCCAATCAC 1684
QY 531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle 550
Db 1685 AACGGTGACCTAGAAAGCGCTCACTGGAATGCAAGGATGCCCAAGAGACTAAGAACCCAGC 1744
QY 551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
Db 1745 GACGACACAATGATGCGCAATAAAT-----GCGGATTTACGCTCCACAGAGAAGCTC 1798
QY 571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db 1799 ACATGAAGATTCAGAGAACCGCAAGGATAGATCAGATGGTGGCTAC-----CCT 1852
QY 591 LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db 1853 TTCATGGGTAATTTCCGGCAATACCAATGGATGAGATGTCAAGATTGTAGTACTCA 1912
QY 607 AspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLeuThrLeu 624
Db 1913 GACCAGGAGCTCATGCGCAAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1972
QY 625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln-----GlnAsnTyr 641
Db 1973 GGTTCACCTCATTTGATGATGTTGTCCTCCAGCACCATCAGGGTTTCATGACAGACCCAC 2032
QY 642 LeuSerAsnAspLeuGlySerArgSerGluMetGly-----SerHis 655
Db 2033 CATGGGATTCCTATAGGAGAGAGAGTGAATAGGAGAAACACAGGAATATGGACCGCC 2092
QY 654 -----SerHis 655
Db 2093 ACCATCAATGGTGTAGTCTCGACCAACCGCACCATTTCATCAGCGGAGCTGCGCGGCT 2152
QY 656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675
Db 2153 TACATGGGATGACATACAGAAC-----CAGAAGAGATATGTGGCT 2194
QY 676 GlnLeuLeuProAspPheVal 682
Db 2195 CAGTTATTCGCCGACTTCGTT 2215
RESULT 5
ADQ37104
ID ADQ37104 standard; DNA; 1929 BP.
XX
AC ADQ37104;
XX
DT 07-OCT-2004 (first entry)
XX
DE Cell proliferation-related nucleic acid sequence #132.
XX
KW cell proliferation related polypeptide; cell proliferation; senescence;
XX
XX differentiation; stress response; ds.
OS Oryza sativa.
XX
XX WO2004061122-A2.
XX
PD 22-JUL-2004.
XX
PF 23-DEC-2003; 2003WO-US041200.
XX
PR 26-DEC-2002; 2002US-0436565P.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
```

PI Cooper B;
XX WPI; 2004-534388/51.
XX New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.
XX
XX Claim 57; SEQ ID NO 263; 408pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related nucleic acid
CC sequence. The present sequence is published separately from the main body
CC of the specification as EPO data.
XX
SQ Sequence 1929 BP; 575 A; 459 C; 463 G; 432 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,03e-67 Length: 1929
Score: 926.50 Matches: 262
Percent Similarity: 51.52% Conservative: 110
Best Local Similarity: 36.29% Mismatches: 225
Query Match: 26.00% Indels: 125
DB: 12 Gaps: 31

US-10-624-201A-2 (1-688) x ADQ37104 (1-1929)
QY 2 TyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHisAsn 21
Db 10 TACTACTCAGCCCTGGCAATGAAAGGACTCGCAAGCTATGATGCCACGCG-----GAT 63
QY 22 HisGlyAsnSerAsnAsnAsnIleGlnThrLeu---TyrLeuMetAsnProAsnAsn 40
Db 64 TCAGGCAATTCATCATATCTGTGCCATCAGCAATAGGAACATGTTATATCTTGGCAAT 123
QY 41 TyrMetGlnGly---TyrThr-----ThrSerAspThrGlnGlnGlnGlnGlnLeu 57
Db 124 GGGTCTCTTGGGCAATACAGCAATTCAGTGGCAATTCATCAGCATCAGCAGAAATTTTCATG 183
QY 58 PheLeuAsnSerSerProAlaAla-----SerAsnAlaLeuCyHisAlaAsnIle 74
Db 184 GAGCTCCCTGGCCATCAACTGGATCTCTCAAGATTCATGTCACGGAACTTAACATG 243
QY 75 GlnHisAlaProLeuGlnGlnHisPheValGlyValProLeuProAlaValSerLeu 94
Db 244 GTCCGCTCGTACATGATCAGCGCTCTTTTGGG-----CTGCCAAAGATATG 291
QY 95 HisAspGlnIleAsnHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSer 114
Db 292 AGAAATGAG-----ATGTTGATGCATCTGATGGATGGAGCACAATATGCT 336
QY 115 GlnGlnValIleValProSerThrGlyValSerAlaThrSerCysGlyGlyIleThr 134
Db 337 GGTGCTGATCTCATCCACATGACATCATAGCAGCGCGAGATTGAGTTTGGCCTATTG 396
QY 135 ThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln 154
Db 397 AACCAACCACAATTCGATGAGCGTT-----GCACCAGCACCGGCCAA----- 438
QY 155 GlnGlnGlnGlnGlnGlyLeuSerLeuSerLeuSerProGlnLeu----- 170
Db 439 -----GGATGTCTCTGAGCCTCAACACGCATATCTCGCGCCTTCG 480
QY 171 -----GlnGlnGlnIleSerPheAsnAsnIleSerSerSer 184
Db 481 TATCCATAGTGTCTCGGAAACAGAGTTGCTAACACCACTCTTACCATGGTGTATGAC 540
QY 185 ProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySerSerSerAsnMetVal 204

Db 541 AACAGAAATGAAGAAT-----ATGCAATCTGAGGCTCACAGGCAATC 582
QY 205 LeuGlySerLysTyrIleuLysAlaAlaGlnGluLeuLeuAspGluValValAsnIleVal 224
Db 583 AGAAACTCAAGTATCTGAAGCAGCACAAGAAATGCTTTGATGAGGTCTCAGT---GTT 639
QY 225 GlyLysSerIleLysGlyAspAspGlnLysLysAsp-----AsnSerMetAsn 240
Db 640 TGGAGAGTATTAAGCAGAGGCTCAGAAAGACCAGGCTGAAGCAGGAAATCAGATAAC 699
QY 241 LysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGluSer 260
Db 700 AAAGAAGCC-----GAGGGGGGTTGCAAAAGTTCAGGGGGTATCT 738
QY 261 SerSerArgGlnLysAsn-----GluValAlaValGluLeuThrAlaGlnArgGln 278
Db 739 TCCAAACCCACAGAGTCTACTGCCAATGCTGCCAGAGATTCTGCTCTCAGAAACA 798
QY 279 GluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluArgTyr 298
Db 799 GAGCTCCAGATTAAGATGCAAAACTTATGGCCATGTTGGATGAGGTGACCGAAGTAC 858
QY 299 ArgGlnTyrHisHisGlnMetGlnIleValLeuSerPheGlnGlnValAlaGlyIle 318
Db 859 AAACACTATTATCATCAAAATGCAATTTAGTCTCATCTTTTGATATGTTGCTGGGTCT 918
QY 319 GlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCys 338
Db 919 GGAGCTGCCAAGCCTTATCTACGTCAGTGGCCCTTCAGACAACTTCAAAACATTTCA 978
QY 339 LeuLysAspAlaIleAlaGlnValLysAlaThrSerLysSerLeuGlyGluGlu 358
Db 979 CTGAAGATGCTATCAACGATCAGATCATGTTATCCGGAAGAAACTTGGAGAGGAG 1038
QY 359 GlyLeuGlyGlyLysIleGluGly-----SerArgLeuLysPheValAspHisLeu 376
Db 1039 AGTTTCATCTGGCAAA---GAGGCAAAATTAACGCGCTCCGTATATATTGACCAATTA 1095
QY 377 ArgGlnArgAlaLeuGlnGlnIleGlyMetGlnProAsnAlaTrpArgProGln 396
Db 1096 AGACAACAGCGCGCTTTCACGAGATGTTGTTGTACAGCAAAATGCTTGGAGGCCACAG 1155
QY 397 ArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeu 416
Db 1156 AGGGGACTGCCCGAAACTCAGTTTCAATTTCTGCTGCTTGGCTTTTGAACATTCCTT 1215
QY 417 HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArg 436
Db 1216 CACCGGTATCCAAAAGATTCAAAAAGCTGATGCTAGCGAGACAACTGGCTTAAACAAG 1275
QY 437 SerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGlu 456
Db 1276 AGTCAGATTCAAAATGGTTTCAATAATGCCGTGCGCTGTGGAACCGATGATCGAA 1335
QY 457 GluMetTyrLeuGluGluValLysAsnGlnGluAsnSerThrAsnThrSerGlyAsp 476
Db 1336 GACATGTATAAGAGAGAGATTGGGGGCGGATCTCGACTCG---AACTCTCTCCCGAC 1392
QY 477 AsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHisProIleThr 496
Db 1393 AACGTACCAAGAGCAGCAAGCAAAATAGCAACATCTGAAGATAAGGAAGT---CTGAAA 1449
QY 497 SerSerLeuLeuGln-----AspGlyIleThrThrGlnAlaGluIleSer 512
Db 1450 AGCTCTATGAGCAGCATTTATCAACAGCCCAATTAGGTGAATCCAAACCCCAACATCGGG 1509
QY 513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHisAlaHisAsn--- 531
Db 1510 ATGATGAGCGCTTGGTGGGCACCA-----GCCGGCTTCCACAACGAGCAACACAG 1560
QY 532 ----PheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle 550

```
Db 1561 GATGACAGCTTCATGAACTTAATG----- 1584
Qy 551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
Db 1585 -----CTGAAGGACCAAGG-----CGGGCGAAGCCGAA 1614
Qy 571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db 1615 GGCAGCCTCTCCATGATGCGTGGCCCATCAATCCGATGAG-----AAGCCT 1662
Qy 591 LeuMetAlaLysAlaMetGlyAspPheGlyArgPheAspProHisAspGlnMet 610
Db 1663 CGGTTATGCTTACCAATTTGCGGGCTCGAAGATAC----- 1701
Qy 611 ThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuPro---ProSer 629
Db 1702 -----GGGACAGCAATGTGTCATTGACACTTGGCTTACAGCATCTGAC 1746
Qy 630 GluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeuGlySerArg 649
Db 1747 AACAGGCTTTCCGTA-----CAGAACACTCATCAGCCAGGTTTCGCGGCTGCTGGA 1797
Qy 650 SerGluMetGlySerHisTyrAsnArgMetGlyTyr-----GluAsnIleAspPhe 666
Db 1798 GAGAAATTTACAACTCCACCGCTTCTCGGTGCTCGCTCGAGCTTCTCTCGGACTAC 1857
Qy 667 GlnSerGlyAsn-----LysArgPhe---ProThrGlnLeuLeuProAsp 680
Db 1858 GAATCCAGCAACCAATAGATCAAGGCAACGGTTCCGAACCATCGCTCTTAATGCATGAT 1917
Qy 681 PheVal 682
Db 1918 TTTGTG 1923

RESULT 6
AD062706
ID AD062706 standard; DNA; 1971 BP.
XX
AC AD062706;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2550 orthologous sequence, SEQ ID 1173.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Glycine max.
XX
XX WO2004031349-A2.
XX
XX 15-APR-2004.
XX
XX 18-SEP-2003; 2003WO-US030292.
XX
XX 18-SEP-2002; 2002US-0411837P.
XX
XX 17-DEC-2002; 2002US-0434166P.
XX
XX 24-APR-2003; 2003US-0465809P.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
XX Riemann JL, Haake V, Dubell AN, Keddie JS, Sherman BK;
XX WPI; 2004-330163/30.
XX
XX New recombinant polynucleotide encoding transcription factor
XX polypeptides, useful for producing transgenic plants with advantageous
XX properties compared to a reference plant.
XX
XX Claim 1; SEQ ID NO 1173; 510pp; English.
```

```
XX
CC The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1971 BP; 667 A; 360 C; 451 G; 493 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1.16e-58 Length: 1971
Score: 826.00 Matches: 216
Percent Similarity: 52.88% Conservative: 78
Best Local Similarity: 38.85% Mismatches: 148
Query Match: 23.18% Indels: 114
DB: 12 Gaps: 18
```

US-10-624-201A-2 (1-688) x ADO62706 (1-1971)

```
Qy 119 ValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAla 138
Db 19 GTCCCAATTTAAAGGGAATTCATCT-----CTTAAGATGATGAGGCC 63
Qy 139 SerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGln 158
Db 64 AGC-----TTACAAAGGAATTGAGAAATGCTGAGTGATGCGCATCTTTAGCTTCT 114
Qy 159 GlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnGlnGlnGln 178
Db 115 TCTGTTGGC-----TTTCAAAA 132
Qy 179 AsnIleSerSerSerProArgThrAsnAsnValThrIle---ArgGlyThrLeuAsp 197
Db 133 AGGGATGGTTGTATATATCCAAACACCCCTCAATGTGCTTGGTGAAGCCAAAGTCAT 192
Qy 198 GlySer-----SerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGlu 215
Db 193 GGATCACAGGCTCGAGCAACAATGCTTGAACCTCAATACCTCAAGGCAGCAGGAG 252
Qy 216 LeuLeuAspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLys 235
```

253 TTGCTTGTGAATAAGTAAAT---GTCCGAAAGGCTTTGAAGCAAACTGGTTTGGAAAAG 309
236 AspAsnSerMetAsnLysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSer 255
310 CAACAGAGTTTCGGTGACACTGGTTTA-----GATGGCTCCAAAGATTCTGAT 357
256 GlyGlyGlyGluSerSerArgGlnLysAsn-----Glu 267
358 GGAATATCTACAGCCCAATCTGGCAGATGCTCTCAGGCCCAATGGTTCGGTGTCTAAC 417
268 ValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuMetLysLysAlaLysLeu 287
418 GCTTCTGTGTGAGCTATCACCTGCAGAACGCGCAGAACTTTGTTGACAGAAGACAAAGCTT 477
288 LeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisGlnMetGlnIle 307
478 TTGTCCATGCTGATGAGGTGGTAAAGATACAGACAGTACTGCCATCAGATGCGAGATT 537
308 IleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeu 327
538 GTGGTGTCTCTTTGACATGGTTGCTGTGGAGCAGCAGAACCATATACACACTT 597
328 AlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnVal 347
598 GCCTTAAGAACAAATTTCTGCCACTTTCCGCTGTTGCTGATGCCATCAGTGGCCAAAT 657
348 LysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyLysIleGluGlySer 367
658 CAGTGACCCAAAGAGCCCTTGGGAGCAAGAGGAATA-----CCC 699
368 ArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMet 387
700 CGTCTCGCTATGTGGATCAGCAACTTAGACAACAAAGGCCCTTCAGCAACTGTGTGA 759
388 MetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeu 407
760 ATGAGA---CAAGCTTGGAGGCCCTCAGAGGGGACTTCCTGAAACCTCTGTTTCAGTACTC 816
408 ArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMet 427
817 CGTGTGTGGCTCTTTGAGCAATTTCCCTTCATCCTTATCCTTAAGGATTCAGAGAAAATATG 876
428 LeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArg 447
877 CTAGCAAGGCAAACTGGCTTAAACAAGAACACAGGTGGCAAACTGGTTCATTATGCAAGG 936
448 ValArgLeuTrpLysProMetValGluGluMetTyrLeuGluValLysAsnGlnGlu 467
937 GTGCGTCTATGGAAGCCAAATGGTTGAGGAAATGTACAAAGAAGAAATTTGGTGTCTGAG 996
468 GlnAsnSerThrAsnThrSer----- 474
997 ATGAGCAGCAATCTATTATCATCAGAGAAACAACTCAAGCTCCAGAGATGATGTTCAA 1056
475 ---GlyAspAsnLysAsnLysGluThrAsn-----IleSerAlaProAsnGluGlu 490
1057 GCTTCTGACATAAAGGGAAGAGTCCAGGACAACTTAATAAACCTGTGATGATGTT 1116
491 LysHis-----ProllelleThrSerSerLeuLeuGlnAspGlylleThrThr 507
1117 CAGCATCATGGGTGAAGTTAGATCATGCCCTCGGAATTTGCACAGAGGGATCCAAAGCAGT 1176
508 -----GlnAlaGlu 510
1177 GATCATGGAGAAAATGCCATGATCTCAGAAATTTGGAATTTGCAAGTGCACCAAGGTTTC 1236
511 IleSerThrSerThrIleSerThrSerProThrAlaGly-----AlaSer 525
1237 AACATGAACAACAGCAATAACATAGTCTCTATTATGGTGTGATGGCTGCATAATGGCTTCT 1296
526 LeuHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThr 545
1297 ACTCTGCCACATATGATTTACAGAGTTAGGTAAACATTTCGCTCGATGCGCCACGTTGTC 1356

QY 546 ThrValAspHisIleGluAsnAsnAlaLysLys-----GlnArgAsnAspMet 561
Db 1357 CTTCGATTGGAATTTAGGAACCTGTGAAAGTCAAGGATTTGGTGTGTCAATGATGACATG 1416
QY 562 HisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArgGlu 581
Db 1417 CATAAACGACATAAGAANAACATTGGCTTCTTCCCGACAGACTGAT----- 1461
QY 582 SerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGly 601
Db 1462 -----TTGCTAGATTACCAT 1476
QY 602 ArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHisGlyAsn 617
Db 1477 TTCACAGACCCAGGAAGCAACAA-----AACAGTTTGGCAAT 1515
RESULT 7
ADO63042
ID ADO63042 standard; DNA; 1617 BP.
XX ADO63042;
AC ADO63042;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2546 coding sequence, SEQ ID 1509.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX Arabidopsis thaliana.
XX
PN WO2004031349-A2.
XX
PD 15-APR-2004.
XX
XX 18-SEP-2003; 2003WO-US030292.
PF
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
XX
DR WPI; 2004-330163/30.
DR P-PSDB; ADO63043.
XX
XX New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
PS Claim 1; SEQ ID NO 1509; 510pp; English.
XX
CC The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance

CC to sugars, altered carbon/nitrogen sensing, early flowering, late
 CC flowering, altered flower structure, loss of flower determinacy, reduced
 CC fertility, altered shoot meristem development, altered branching pattern,
 CC altered stem morphology, altered vascular tissue structure, reduced
 CC apical dominance, altered trichome density, altered trichome development,
 CC altered trichome structure, altered root development, altered shade
 CC avoidance, altered seed development, altered seed ripening, altered seed
 CC germination, slow growth, fast growth, altered cell differentiation,
 CC altered cell proliferation, altered cell expansion, altered phase change,
 CC altered senescence, abnormal embryo development, altered programmed cell
 CC death, lethality when overexpressed, altered necrosis patterns, increased
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark
 CC green leaves, change in leaf shape, increased leaf size and mass, light
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
 CC altered seed coloration, altered seed size, altered seed shape, large
 CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
 CC content, altered seed protein content, altered seedprenyl content,
 CC altered leaf prenyl lipid content, increased anthocyanin levels, and
 CC decreased anthocyanin levels. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 1617 BP; 510 A; 335 C; 358 G; 414 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.18e-58 Length: 1617
 Score: 824.50 Matches: 212
 Percent Similarity: 53.99% Conservative: 79
 Best Local Similarity: 39.33% Mismatches: 111
 Query Match: 23.14% Indels: 137
 DB: 12 Gaps: 21

US-10-624-201A-2 (1-688) x ADO63042 (1-1617)

Qy 109 AsnAsnGlnAspGlnSerGlnGlnValleValProSerSerThrGlyValSerAlaThr 128
 Db 58 AACCATCAACAAACAAACCAACAGCT---TCTTCTTCATCCGCCGCCGCTCTTTCCACC 114
 Qy 129 SerCysGlyGlyIleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIlePro 148
 Db 115 GTACCCGGCGGCGATCT-----GTTCGAAGCAGATGTTTATCCACCACCACC 168
 Qy 149 ThrProGln-----HisArg 153
 Db 169 ACAGGACGCTGTGAACAGAAACGGTACCGTTTCAAGCAGCGATCTAAGCTTTCAGCAT 228
 Qy 154 GlnGlnGlnGlnGlnGlnGlyLeuSerLeuSerLeuSerProGlnLeu----- 170
 Db 229 GGTCAA-----GGACTGCTTTGAGCCTTGGTACTCAGATCTCTGTGTGCT 273
 Qy 171 -----GlnGlnGlnIleSerPheAsnAsn-----AsnIle 180
 Db 274 CCGTTTCACCTTCATCAATACCAATGCGGTTTACTAGTCAGATCCCTCAATTCAGTC 333
 Qy 181 SerSerSerPro-----ArgThrAsnAsnValThrIle 192
 Db 334 AAGGAAACGTCACCGTTTCATGTGGATGAGATGAGTGTGAAGAGCAAGAAATGATCTTG 393
 Qy 193 ArgGlyThrLeuAspGlySerSerSer----- 201
 Db 394 TTGGGTCAATCTGATCTCTCTCTGTTATGCTGGTAAATGGTGGAAATGGCTTCTACAAC 453
 Qy 202 -----AsnMetValLeuGlySerLysTyr 209
 Db 454 AATTATCGGTATATAGACATACAGAGGGTTTATGAGCAGCGTTCTCGGTCTTCGGTAT 513
 Qy 210 LeuLysAlaAlaGlnGluLeuLeuAspGluValValAsnIleValGlyLysSerIleLys 229
 Db 514 CTTAAACCTGCTCAGAAATTTGCTTGAATGAAGTGGTTAGTGTC----- 555
 Qy 230 GlyAspAspGlnLysLysAspAsnSerMetAsnLysSerMetProLeuAlaSerAsp 249
 Db 555

Db 556 -----AAGAAAGAACTAAACCAAAATCGGGAAGAAAGATGAAAGTT---AATGAC 603
 Qy 250 ValAsnThrAsnSer-----SerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 265
 Db 604 TTTAACAGCTGTTCTAAGGAGATAGAGGAGGAGGTGTCAGTTATCGAGT---GATTCTG 660
 Qy 266 AsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGlnLeuGlnMetLysLysAla 285
 Db 661 AATCGGAAATCGATTGAGTTATCTACAATTTGAACGCTGAAGAGCTTCAGAAAGAAAC 720
 Qy 286 LysLeuLeuAlaMetLeuGluValGluGlnArgTyrArgGlnTyrHisHisGlnMet 305
 Db 721 AAGCTTTTAAATGTTGATAGGTAGATAAAGATATAACCAATATTACCAATCAATG 780
 Qy 306 GlnIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThr 325
 Db 781 GAAGCATTAGCTTCATCATTTGAGATAGTAGCAGGACTTGGATCAGCTTAAGCCTTACACA 840
 Qy 326 GlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGlu 345
 Db 841 TCAGTTGCTCTCAACAGAAATCTCTCGCCATTTTCGTCTCTCGCGACGCAATAAAGGAA 900
 Qy 346 GlnValLysAlaThrSerLysSerLeuGlyGlu-----GluGluGly 359
 Db 901 CAAATTCAGATTGTTAGAGAAACCTTGGGGAAGAAAGGAGAGTCTGTTGATGAGCAA 960
 Qy 360 LeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGln 379
 Db 961 CAAGGAGAGAGGATA-----CCAGGTTGAGGTATTAGATCAACCGTTGAGACACGNA 1014
 Qy 380 ArgAlaLeu---GlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGly 398
 Db 1015 AGAGCTTTGCATCAACAGCTTGAATGTTTCGACCT---GCTTGGAGACCTCAAGAGGT 1071
 Qy 399 LeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisPro 418
 Db 1072 CTTCTCGAAACCTCTGTCTCTGTTCTTTCGCGCTTGGCTCTTTGAACATTTCTTCATCCA 1131
 Qy 419 TyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGln 438
 Db 1132 TATCCGAAAGAAATCTGAGAAATCATGCTTTCGAAAGCAGACAGGATTTGTCAAAGAACCCAG 1191
 Qy 439 ValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMet 458
 Db 1192 GTTGTAAATTTGTTTCAAAACGCGAGAGTTCGTCTGTGGAACCGGATGATCGAAGAAATG 1251
 Qy 459 TyrLeuGluGluVal-----LysAsnGlnGluGln 468
 Db 1252 TATAAAGAGAGTTGTTGGTATGAATCTGAGTTACTTAATCTCCAAATCTTCCCAAGAACCC 1311
 Qy 469 AsnSerThrAsnThrSerGlyAspAsn-----LysAsnLys 480
 Db 1312 AACAGCAACAAACCAAGAGACTCTCATCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1371
 Qy 481 GluThrAsnIleSerAlaProAsnGluGlyHisProIleIleThrSerSer----- 498
 Db 1372 AACAGCAACCTCGCTTATTCATCTGCAGACACAAACAACTTGTCTTCTCATCAGAAACC 1431
 Qy 499 -----LeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSerThrSer 514
 Db 1432 AAACAGATCGTGTCTTAGCAATGATTAACGCCACAGCAACCAACAGATAAACCGCTCA 1491
 Qy 515 Thr-----Ile 516
 Db 1492 TCGGATTACGACACTCTGATGAATATATCAGGGTTTGGTGTGTGATGATTACCGTTACATT 1551
 Qy 517 SerThrSer-----ProThrAlaGlyAlaSerLeuHisHisAlaHisAsnPhe 532
 Db 1552 AGCGGAAGCAACAGCAAGAAAGCAGATTTCCAAATTCCTCATCTTACACGACTTT 1608
 RESULT 8
 ABZ13100
 ID ABZ13100 standard; DNA; 1884 BP.


```
QY 432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
Db 1390 ACTGGTTATCCAGAAATCAGGTCAAAATGGTTCTATAATGCTAGGTTTCGCCTATGG 1449
QY 452 LysProMetValGluGluMetTyrLeuGluGluValLys----- 464
Db 1450 AAACCAATGGTGGAGAGATGTATCAACAAGAGCAAAAGAAAGAGAGAGAGAGAA 1509
QY 465 -----AanGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
Db 1510 GAAATGAAATCAACAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1569
QY 481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
Db 1570 CCCAACCAAC-----AATGAAACAACTTCTACTGTATTAACCGCACAA----- 1611
QY 501 GlnAspGlyIleThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
Db 1612 -----ACTCCA 1617
QY 521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
Db 1618 ACGACGATGATCGACACATCAC----- 1641
QY 541 AspAsnThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAsp 560
Db 1642 -----GAAACGAC 1650
QY 561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580
Db 1651 -----TCTTCATTCCTCTCTTCGCTC-----GCCGCCGCTTCT 1683
QY 581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600
Db 1684 CACGGCGGTTTCAGACGGGTTCCACGTCGCCACGCTG----- 1719
QY 601 GlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHis-----Gly 616
Db 1720 -----CAGCAAGACGTCAGTGACCTCCACGTCGCGGAGATGGT 1758
QY 617 AsnAsn-----GlyValSerLeuThrLeuGlyLeu 626
Db 1759 GTGAACGCTATAGATTCCGGGACCAACACAGCTGGTGACGTGCTCTTACGCTTGGCTA 1818
QY 627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
Db 1819 CGCCACTCTGCAATATT-----CCTGATAAGAACACTTCTTCTCCGTTAGAGACTTT 1872
QY 647 Gly 647
Db 1873 GGA 1875
RESULT 9
AAD06494
ID AAD06494 standard; cDNA; 1983 BP.
XX AC AAD06494;
XX AC
XX 10-AUG-2001 (first entry)
XX DE Arabidopsis thaliana G418 transcription factor homologue, G965 cDNA.
XX KW Transcription factor; pesticidal; antimicrobial; gene therapy;
XX KW pathogen tolerance; trichome structure; callose induction;
XX KW phytoalexin induction; plant structure; plant development; ss.
XX OS Arabidopsis thaliana.
XX Key Location/Qualifiers
XX CDS 73..1956
XX FT /*tag= a
XX FT /product= "A. thaliana transcription factor homologue"
XX FT
XX
```

```
PN WO200135726-A1.
XX PD 25-MAY-2001.
XX PF 14-NOV-2000; 2000WO-US031418.
XX PR 17-NOV-1999; 99US-0166228P.
XX PR 17-APR-2000; 2000US-0197899P.
XX PR 22-AUG-2000; 2000US-0227439P.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (HEAR/) HEARD J.
PA (RATC/) RATCLIFFE O.
PA (CREE/) CREELMAN R.
PA (JIAN/) JIANG C.
PA (PINE/) PINEDA O.
PA (REUB/) REUBER L.
PA (ADAM/) ADAM L.
XX
PI Heard J, Ratcliffe O, Creelman R, Jiang C, Pineda O, Reuber L;
PI Adam L;
XX WPI: 2001-335978/35.
XX P-PSDB; AAE02525.
DR
XX Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the pathogen resistance characteristics of plants, e.g.
PT corn, potato and cotton plants.
XX
PS Claim 4; Page 114-116; 134pp; English.
XX
CC The present sequence is a cDNA encoding Arabidopsis thaliana
CC transcription factor homologue. The transcription factors are used to
CC alter the structure and developmental characteristics of plants such as
CC soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,
CC alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry,
CC raspberry, carrot, cantaloupe, cauliflower, cucumber, coffee, eggplant,
CC grapes, mango, lettuce, honeydew, melon, onion, papaya, peppers,
CC pineapple, spinach, squash, sweet corn, tobacco, tomato, peas,
CC watermelon, rosaceous fruits and vegetable brassicas. The transcription
CC factors are specifically useful for modifying traits associated with
CC plant's pathogen tolerance such as alterations in cell wall composition,
CC trichome number or structure, callose induction, phytoalexin induction,
CC and alterations in the cell death response. Transgenic plants expressing
CC these transcription factors are more tolerant to biotrophic or
CC necrotrophic pathogens such as fungi, bacteria, molluscs, viruses,
CC nematodes and parasitic higher plants. The transcription factors are also
CC used in gene therapy
XX
SQ Sequence 1983 BP; 585 A; 464 C; 434 G; 500 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1-72e-58 Length: 1983
Score: 824.00 Matches: 237
Percent Similarity: 45.96% Conservative: 76
Best Local Similarity: 34.80% Mismatches: 178
Query Match: 23.13% Indels: 190
DB: 4 Gaps: 25
US-10-624-201A-2 (1-688) x AAD06494 (1-1983)
QY 15 AspHisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyr 34
Db 331 GAGCAACAAACCGTCAC-----AACCAACAACTAATCACTCACTTTCAT 375
QY 35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrSerAspThrGlnGlnGln 54
Db 376 ATGTTATTACCAATCATCAT---CAAGGTTTGTCTTCCACGACGAAACACTATGAC 432
QY 55 GlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle 74
Db 432 ----- 432
```

```
QY 75 GlnHisAlaProLeuGlnGlnHisPheValGlyValProLeuProAlaValSerLeu 94
Db 433 -----CCGCAGCAACAACACTTTACA-----TGGCCATCTTCCTCC--- 471
QY 95 HisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSer 114
Db 472 -----TCCGATCATCAT-----CAAAACCGAGATATGATC 501
QY 115 GlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThr 134
Db 502 GGAAACCGTCCACGTGGAGAGAAAGGGTTTGTCTTTATCT-----CTCTCA 549
QY 135 ThrAspLeuAlaSerGlnLeuAlaPheGln---ArgProIle----- 147
Db 550 TCTTCATTAGCCGACGTAAGCCGAGGAATATAGAACGATTATTGTGCACGCCGTTGAT 609
QY 148 -----ProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGlnGly 161
Db 610 GGAACTTCTTCTTCTTAACGCATCGGCTCATCATCAATCAATCAATCAATCAAGAAAT 669
QY 162 LeuSerLeu-----SerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsn 179
Db 670 CTTCTTCTTCTGAGAAATCTTCTTCTCAACATCATCACCATCAAGTTGTGGACATTTTGT 729
QY 180 IleSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySer 199
Db 730 TCATCATCATCATCTCCCATGGCGGCTTCTTCATCCATTGGAGGATCTAC----- 780
QY 200 SerSerAsnMetValLeuGlySerIleValLeuGlyAlaGlnGlnLeuLeuAspGlu 219
Db 781 -----ACGTTGAGGAATTCGAAATATACGAAACCGGCTCAAGAGTTGTTGAAGAG 831
QY 220 ValValAsnIleValGlySerIleLeuGlyAspAspGlnIleValAspAsnSerMet 239
Db 832 TTTTGTAGTGT-----GGAGAGGACATTTCAAGAGAACAAACTT 873
QY 240 AsnLysGluSerMetProLeuAlaSerAspValAsnThr-----AsnSerSerGlyGly 257
Db 874 AGTAGGAACAAC-----TCAAAACCTTAATACTACCGGTGGAGGAGGCGGA 921
QY 258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db 922 GGGTCTCTCGTCATCGCGCGGAACAGCTAATGATGATGATGATGATGATGATGATGATG 981
QY 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGlnValGlnArg 297
Db 982 ATTGAACATCAAGAAGAAAGTCAAGCTACTACTATCTTCTTGAAGAGTTGACCCAGG 1041
QY 298 TyrArgGlnTyrHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
Db 1042 TACAACCCTACTCTCGCAACAAATGCAATGGAATGGTAGTGAACCTCATTCGACCAAGTAATGGT 1101
QY 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db 1102 TACGCGCGCGCGGTTCGTTACACGACATAGCTCAAAAGGCAATGTCTAGGCATTTCCGG 1161
QY 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 1162 TGTTTGAAGACCGGTAGCGGTTCAGCTTAACGCGAGCTGTGAGCTTCTAGGGGATATA 1221
QY 358 GluGlyLeuGlyGly-----LysIleGluGlySerArgLeuLysPhe 371
Db 1222 GAGCGCGAGGGCTGCATCTCGGGTTAAACCAAGGGGAAACGCGCGATTCGCTTGG 1281
QY 372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
Db 1282 CTAGACGAGAGTTTGGCTCAGCAACAGCGGTTTCATCATATGGGTATGATGGAGCAAGAG 1341
QY 392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuAlaTrpLeu 411
Db 1342 GCATGAGACCGCAAGCTGTTTGGCTTGAACGCTCCGTTAATATCTTAGAGCTTGCTA 1401
QY 412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
```

```
Db 1402 TTCGAGCATTTTCTTAATCCGTACCAAGCGATGCTGATAAGCACCTCTTAGCACGACAG 1461
QY 432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
Db 1462 ACTGGTTTATCCAGAAATCAGGTGTCAAAATGGTTTCATAAATGCTAGGGTTGCGCTTGG 1521
QY 452 LysProMetValGluGluMetTyrLeuGluGluValLys----- 464
Db 1522 AAACCAATGTTGGAAGAGATGTATCAACAAGAAAGCAAAAGAAAGAAAGCAAGAGAA 1581
QY 465 -----AsnGlnGlnGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
Db 1582 GAAATGAAATCAACAACAAGACAGACAGCAACAACAACAACAACAACAACGACAGAAA 1641
QY 481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
Db 1642 CCCAACAAAC-----AATGAAACAACACTTCACTGTCTATACCCGCAAA----- 1683
QY 501 GlnAspGlyIleThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
Db 1684 ----- 1689
QY 521 ThrAlaGlyAlaSerLeuHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
Db 1690 ACGACGATGACATCGACACATCAC----- 1713
QY 541 AspAsnThrThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAsp 560
Db 1714 -----GAAACGAC 1722
QY 561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580
Db 1723 -----TCCTCATCTCTCTTCCGTC-----GCCGCCGCTTCT 1755
QY 581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600
Db 1756 CACGGCGGTTACAGCGGTTCCCGTCGCGCATGT----- 1791
QY 601 GlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHis-----Gly 616
Db 1792 -----CAGCAAGACGTCAGTGACTTCCACGTCGACGAGATGGT 1830
QY 617 AsnAsn-----GlyValSerLeuThrLeuGlyLeu 626
Db 1831 GTGAACGTCATAAGATTCCGGACCAACAGACGTTGGTGAAGTGTCTCTTACGCTTGTCTA 1890
QY 627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
Db 1891 CGCCACTCTGCAATATT-----CCTGATAAGAACACTTCTTCTCCGTTAGAGACTTT 1944
QY 647 Gly 647
Db 1945 GGA 1947
RESULT 10
ADC46618
ID ADC46618 standard; DNA; 1983 BP.
XX
AC ADC46618;
XX
DT 18-DEC-2003 (first entry)
XX
DE Thalecress transcription factor-like DNA G965.
XX
KW Thalecress; transcription factor-like protein; ds; seed trait;
KW transgenic; gene; plant size; stress tolerance; yield;
KW disease resistance; plant.
XX
OS Arabidopsis thaliana.
XX
FN US2003093837-A1.
XX
```


PD 15-MAY-2003.
 XX 01-NOV-2002; 2002US-00286264.
 XX 23-MAR-1999; 99US-0125814P.
 XX 22-MAR-2000; 2000US-00533030.
 XX (KEDD/) KEDDIE J.
 PA (RIEC/) RIECHMANN J L.
 PA (RATC/) RATCLIFFE O.
 PA (ZHAN/) ZHANG J.
 PA (JIAN/) JIANG C.
 PA (PINE/) PINEDA O.
 PA (HEAR/) HEARD J.
 PA (YUGG/) YU G.
 PA (ADAM/) ADAM L.
 PA (BROU/) BROUN P.
 PA (REUB/) REUBER L.
 PA (FILG/) FILGRIM M.
 PA (SAMA/) SAMAHA R.
 XX
 PI Keddle J, Riechmann JL, Ratcliffe O, Zhang J, Jiang C, Pineda O;
 PI Heard J, Yu G, Adam L, Broun P, Reuber L, Pilgrim M, Samaha R;
 XX
 DR WPI; 2003-765498/72.
 DR P-PSDB; ADC46619.
 XX
 PT Novel transgenic plant having recombinant polynucleotide encoding
 PT polypeptide that alters trait of transgenic plant when compared with same
 PT trait of another plant lacking recombinant polynucleotide.
 XX
 PS Disclosure; SEQ ID NO 17; 165pp; English.
 XX
 CC The invention relates to a transgenic plant having recombinant
 CC polynucleotide (II) encoding polypeptide comprising at least 6
 CC consecutive amino acids of a sequence chosen from the protein sequence
 CC appearing as ADC46603 - ADC46749 (every second sequence), where
 CC recombinant polynucleotide alters a trait of the seed transgenic plant
 CC when compared with same trait of another plant lacking recombinant
 CC polynucleotide. The proteins are transcription factor-like proteins. Also
 CC included are altering (M1) a trait associated with seed (comprising:
 CC transforming a plant with (II); selecting the transformed plants; and
 CC identifying a transformed plant with seed having altered trait), altering
 CC (M2) the expression levels of at least one gene of a plant (involving
 CC transforming the plant with (II) and selecting the transformed plant),
 CC altering (M3) a trait associated with a plant's seed (comprising:
 CC transforming the plant with a recombinant polynucleotide comprising a
 CC nucleotide sequence comprising least 18 consecutive nucleotides of a
 CC sequence appearing as ADC46750 - ADC46766 and selecting the transformed
 CC plan) altering (M4) a plant's trait (involving providing a database
 CC sequence, comparing the database sequence with a polypeptide or a
 CC polynucleotide chosen as detailed above, selecting a database sequence
 CC that needs selected sequence criteria and transforming a database
 CC sequence in the plant) and altering a plant's trait (involving providing
 CC a test polynucleotide, hybridising the test polynucleotide with a
 CC polynucleotide detailed above at low stringency and transforming the
 CC hybridising test polynucleotide in a plant to alter a trait of the
 CC plant). The method (M1) is useful for altering a trait associated with
 CC seed. The method (M2) is useful for altering the expressing levels of at
 CC least one gene of a plant. The method (M3) is useful for altering a trait
 CC associated with a plant's seed. The method (M4) is useful for altering a
 CC plant's trait. The method (M4) is useful for altering a plant's trait
 CC such as seed or plant size, stress tolerance, yield or disease
 CC resistance. The present sequence encodes a transcription factor-like
 CC protein/seed trait altering protein of the invention.
 XX
 SQ Sequence 1983 BP; 586 A; 464 C; 433 G; 500 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.72e-58 Length: 1983
 Score: 824.00 Matches: 237
 Percent Similarity: 45.96% Conservativeness: 76
 Best Local Similarity: 34.80% Mismatches: 178

Query Match: 23.13% Indels: 190
 DB: 10 Gaps: 25
 US-10-624-201A-2 (1-688) x ADC46618 (1-1983)
 QY 15 AspHisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnAsnGlnThrLeuTyr 34
 DB 331 GAGCAACAACACCGTCAC-----AACAAACAATAAATCACTCACTTTCAT 375
 QY 35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGln 54
 DB 376 ATGTTATTACCAATCATCAT---CAAGGTTTGTCTTCCACGAGAAACACTATCCAG 432
 QY 55 GlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsn 74
 DB 432 ----- 432
 QY 75 GlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeu 94
 DB 433 -----CCGACGACCAACAACACTTTTACA-----TGGCCATCTTCCTCC--- 471
 QY 95 HisAspGlnIleAsnHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSer 114
 DB 472 -----TCCGATCATCAT-----CMAAACCGAGATATGATC 501
 QY 115 GlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThr 134
 DB 502 GGAACCGTCCACGTGAAGGAGGAAAGGGTTGTCTTTATCT-----CTCTCA 549
 QY 135 ThrAspIleAsnSerGlnIleAlaPheGln---ArgProIle----- 147
 DB 550 TCTTCATTAGCCGACGTAAGCCGAGGAGATATAGAAGCATTTATTGTGCGACGGTTGAT 609
 QY 148 -----ProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGly 161
 DB 610 GGAACCTTCTTCTTCTTAACGCATCGCTCATCATCATCATCATCATCATCAAGTTCAAGAAT 669
 QY 162 LeuSerLeu-----SerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsn 179
 DB 670 CTCTCTCTTGAGAAATCTCTCTCTCAACATCATCATCATCAAGTTGTGGACATTTTGGT 729
 QY 180 IleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySer 199
 DB 730 TCATCATCATCATCTCCCATCGCGGTTCTTCATCATCATCGAGGATGTTGGAAGAG 831
 QY 200 SerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGlnLeuLeuAspGlu 219
 DB 781 -----ACGTTGAGGAATTCGAATATACGAACCGCTCAAGATTGTTGGAAGAG 831
 QY 220 ValValAsnIleValGlyLysSerIleLysGlyAspGlnLysLysAspAsnSerMet 239
 DB 832 TTTTGTGTGTT-----GGAAGAGGACATTTCAAGAGAGAACAAACTT 873
 QY 240 AsnLysGluSerMetProLeuAlaSerAspValAsnThr-----AsnSerSerGlyGly 257
 DB 874 AGTAGGAACAAC-----TCAACCCCTAATACTACCGGTGGAGGAGGAGCGGA 921
 QY 258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrAlaGlnArg 277
 DB 922 GGTCTCTGTCATCGCGCGGACAGCTAATGATGATGCTCTCTTGTCTCGGCTGATCGG 981
 QY 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
 DB 982 ATTGAACATCAAGAAGAAAGCAAGCTACTATCTATGTTGTAAGAGGTGGACCGACGG 1041
 QY 298 TyrArgGlnTyrHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
 DB 1042 TACAACCACTACTCGAACAATAATGCAATGGTAGTGAACCTTCGACCAAGTAATGGGT 1101
 QY 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
 DB 1102 TACGGCGCGCGGTTCGTACACGACATAGCTCAAAAGCAATGCTAGGCATTTCCGG 1161

```
QY 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 1162 TGTTCGAAGACGCGGTAGCGGTTACGTTAAACGACGCTGTGAGCTTCTAGGGGATAAA 1221
QY 358 GluGlyLeuGlyGly-----LysIleGluGlySerArgLeuLysPhe 371
Db 1222 GAGCGCGCAGGGGCTGCATCCTCGGGGTTAACCAAGAGGGGMAACGCGCGATTCGCTTG 1281
QY 372 ValAspHisHisLeuArgGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
Db 1282 CTAGACAGAGTTTGGCTGAGCAACAGCGCTTTCATCATATGGTATGATGGAGCAAGAG 1341
QY 392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
Db 1342 GCATGGAGACCGCAACGTTGGTTCCTGAACGCTCGCTTAATATCTTAGAGCTTGCTTA 1401
QY 412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
Db 1402 TTCGAGCATTTTCTTAATCCGTACCCAAGCGATGCTGATAAGCACCTCTTAGCAGCACAG 1461
QY 432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
Db 1462 ACTGGTTTATCCAGAATCAGGTGTCAAATTTGGTTTCATAATGCTAGGGTTGCGCTATGG 1521
QY 452 LysProMetValGluGluMetTyrLeuGluGluValLys----- 464
Db 1522 AAACCAATGGTGAAGAGATGTATCAACAGAGCAAAAGAAAGAGAGAGCAGAGAA 1581
QY 465 -----AsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
Db 1582 GAAATATGAAATATCAACAACAACAAGAGACAGCAACAACAACAACAACACGACAGAAA 1641
QY 481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
Db 1642 CCCAACAC-----AATGAAAACAACCTTCAGTGTCAATACCGCACAA----- 1683
QY 501 GlnAspGlyIleThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
Db 1684 -----ACTCCA 1689
QY 521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
Db 1690 ACGACGATGACATCGACACATCAC----- 1713
QY 541 AspAsnThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAsp 560
Db 1714 -----GAAAACGAC 1722
QY 561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580
Db 1723 -----TCCTCATTTCTCTCTTCGTC-----GCCGCGCTTCT 1755
QY 581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600
Db 1756 CACGCGCGTTCAGACCGGTTCCACCGTCGCCACGCTGT----- 1791
QY 601 GlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHis-----Gly 616
Db 1792 -----CAGCAAGACGTCGATGCTTCACGTCGACGAGATGGT 1830
QY 617 AsnAsn-----GlyValSerLeuThrLeuGlyLeu 626
Db 1831 GTGAACGTCATAAGATTCCGGGACCAACACAGACTGGTGACGTGCTCTTAGCCTTGCTTA 1890
QY 627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
Db 1891 CGCCACTCTGGCAATATT-----CCTGATAGAACACTTCTTCTTCCTAGAGACTTT 1944
QY 647 Gly 647
Db 1945 GGA 1947
RESULT 11
```

```
ADD30786
ID ADD30786 standard; cDNA; 1983 BP.
AC ADD30786;
XX
DT 15-JAN-2004 (first entry)
XX
DE Plant yield-related polynucleotide clone G965.
XX
ds; transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.
XX
OS Arabidopsis thaliana.
XX
PN WO2003013227-A2.
XX
PD 20-FEB-2003.
XX
PF 09-AUG-2002; 2002WO-US025805.
XX
PR 09-AUG-2001; 2001US-0310847P.
PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-0338692P.
PR 14-JUN-2002; 2002US-00171468.
XX
(MEND-) MENDEL BIOTECHNOLOGY INC.
XX
Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
PI Broun PE;
XX
DR WPI: 2003-248221/24.
DR P-PSDB; ADD30787.
XX
PT New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
PS Disclosure; SEQ ID NO 815; 454pp; English.
XX
The invention relates to a number of isolated Arabidopsis thaliana cDNA
sequences and their encoded proteins which are especially transcription
factor related cDNA's and proteins. The isolated or recombinant plant
transcription factor polynucleotides and polypeptides are useful in
producing transgenic plants with commercially valuable properties, i.e.
modified or altered desirable traits as compared to a reference plant,
such as an alteration in a plant growth characteristic, e.g. growth rate,
germination rate of seeds, vigor of plants and seedlings, or leaf and
flower senescence. Sequence information related to the polynucleotides
and polypeptides can also be used in bioinformatic search methods. The
transgenic plant is useful for growing a progeny plant from a parent
plant. This sequence represents one of the cDNAs of the invention.
XX
SQ Sequence 1983 BP; 585 A; 464 C; 434 G; 500 T; 0 U; 0 Other;
Alignment Scores:
pred. No.: 1-72e-58 Length: 1983
Score: 824.00 Matches: 237
Percent Similarity: 45.96% Conservative: 76
Best Local Similarity: 34.80% Mismatches: 178
Query Match: 23.13% Indels: 190
DB: 10 Gaps: 25
US-10-624-201a-2 (1-688) x ADD30786 (1-1983)
QY 15 AspHisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnIleGlnThrLeuTyr 34
Db 331 GAGCAACAACAAACCGTCAC-----AACCAACAACAATACTCACTCTTCAT 375
QY 35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGln 54
Db 376 ATGTTATTACCAATCATCAT---CAAGGTTTTGCTTCCACCGCAAAACACTATGTCAG 432
```


branching pattern; apical dominance; trichome; stem morphology;
 KW root growth; root hair; seed development; cell proliferation;
 KW cell differentiation; premature senescence; necrosis; plant size;
 KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
 KW plant anthocyanin; light response; shade avoidance; bioinformatic;
 KW transcription factor; ds.
 XX Unidentified.

XX US2004019927-A1.

XX 29-JAN-2004.

XX 25-FEB-2003; 2003US-00374780.

XX 18-APR-2001; 2001US-00837944.

XX (SHER/) SHERMAN B K.
 PA (RIEC/) RIECHMANN J L.

PA (JIAN/) JIANG C.

PA (HEAR/) HEARD J E.

PA (HAAK/) HAAKE V.

PA (CREE/) CREELMAN R A.

PA (RATC/) RATCLIFFE O.

PA (ADAM/) ADAM L J.

PA (REUB/) REUBER T L.

PA (KEDD/) KEDDIE J.

PA (BROU/) BROWN P E.

PA (PILG/) PILGRIM M L.

PA (DUBE/) DUBELL A N.

PA (PINE/) PINEDA O.

PA (YUGG/) YU G.

XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;

PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddle J, Brown PE;

PI Pilgrim ML, Dubell AN, Pineda O, Yu G;

XX WPI; 2004-132245/13.

DR P-PSDB; ADI44019.

XX New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.

XX Disclosure; SEQ ID NO 2481; 435pp; English.

CC The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produce a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure, change in stem bifurcations, altered branching
 CC pattern, reduced apical dominance, reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC altered cell proliferation or cell differentiation; rapid development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins, or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This sequence represents a plant
 CC transcription factor related polynucleotide.

XX SQ Sequence 1983 BP; 585 A; 464 C; 434 G; 500 T; 0 U; 0 Other;

Alignment Scores:

Prod. No.: 1-72e-58 Length: 1983

Score: 824.00 Matches: 237

Percent Similarity: 45.96% Conservative: 76

Best Local Similarity: 34.80% Mismatches: 178

Query Match:	23.13%	Indels:	190
DB:	12	Gaps:	25
US-10-624-201A-2 (1-688) x ADI44018 (1-1983)			
QY 15 AspHisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnGlnGlnThrLeuTyr 34			
DB 331 GAGCAACAAAACCGTCAC-----AACCAACAACAATAACTCAACTCTTCAT 375			
QY 35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrSerAspThrGlnGlnGln 54			
DB 376 ATGTTATTACCAATCATCAT---CAAGGTTTTGCTTCCACCGACAAAACACTATGCAG 432			
QY 55 GlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle 74			
DB 432 -----			
QY 75 GlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeu 94			
DB 433 -----CCGACGACCAACAACACTTTTACA-----TGGCCATCTTCTCTCC--- 471			
QY 95 HisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSer 114			
DB 472 -----TCCGATCATCAT-----CAAAACCGAGATATGATC 501			
QY 115 GlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThr 134			
DB 502 GGAACCGTCCACGTGAAGAGGAAAGGTTGTCTTTATCT-----CTCTCA 549			
QY 135 ThrAspLeuAlaSerGlnLeuAlaPheGln---ArgProIle----- 147			
DB 550 TCTTCATTAGCCGACGCTAAAGCCGAGGAATATAGAAGCATTTATTGTGCAGCCGTTGAT 609			
QY 148 -----ProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGlnGly 161			
DB 610 GGAACCTTCTTCTTCTTAAACGATCCGCTCATCATCAATTCATCAGTTCAAGAAT 669			
QY 162 LeuSerLeu-----SerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsn 179			
DB 670 CTCTCTCTGAGAAATCTTCTCTCAACATCATCACCATCAAGTTGTGGACATTTTGGT 729			
QY 180 IleSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySer 199			
DB 730 TCATCATCATCATCTCCATGGCGCTTCTTCTCATCCATTCGAGGATCTAC----- 780			
QY 200 SerSerAsnMetValLeuGlySerIleValLeuAlaAlaGlnGlnLeuLeuAspGlu 219			
DB 781 -----ACGTGAGGAATTCGAAATATACGAACCGGCTCAAGAGTTGTGGAAGAG 831			
QY 220 ValValAsnIleValGlySerIleValGlyAspAspGlnIleValAspAsnSerMet 239			
DB 832 TTTTGTAGTGT-----GGAGAGGACATTTCAAGAGAACAACTT 873			
QY 240 AsnLysGluSerMetProLeuAlaSerAspValAsnThr-----AsnSerSerGlyGly 257			
DB 874 AGTAGGAACAAC-----TCAAAACCCCTAATACTACCGGTGGAGGAGGCGGA 921			
QY 258 GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277			
DB 922 GGGTCTCTCTATCGCGCGGAACAGTAATGATAGTCCTCTTGTCTCCGGCTGATCGG 981			
QY 278 GlnGluGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGlnArg 297			
DB 982 ATTGAACATCAAGAAGAAAGTCAAGCTACTATCTATCTTGAAGAGGTGGACCGCG 1041			
QY 298 TyrArgGlnTyrHisGlnMetGlnIleValLeuSerPheGluGlnValAlaGly 317			
DB 1042 TACAACCACTACTCGCAACAATAATGCAATGCTAGTGAACATCTTCACCAAGTAATGGGT 1101			
QY 318 IleGlySerAlaLysSerThrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337			
DB 1102 TAGCGCGCGCGGTTCCGTACACGACATAGCTCAAAAGGCAATGCTAGGCAATTTCCGG 1161			

Qy 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 1162 TGTGTTGAAAGACGGGTAGCGGTTTCAAGCTTAAACGACGCTGTAGGCGATAAA 1221
Qy 358 GluGlyLeuGlyGly-----LysIleGluGlySerArgLeuLysPhe 371
Db 1222 GAGCGCGGAGGGGTGCATCTCGGGGTTAAACAAAGGAGAAACGCGCGATTTCGGTTG 1281
Qy 372 ValAspHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
Db 1282 CTAGACGAGATTTCGTCAGCAACAGCGGTTTCATCATATGGGTATGATGGACAGAG 1341
Qy 392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
Db 1342 GCATGGAGACGCAACGCGTTCCTGCTGACGCTCCGTTAATATCTTAGAGCTTGGCTA 1401
Qy 412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
Db 1402 TTCGAGCATTTTCTTAATCCGTACCCAGCGATCTGATAGCACCTCTTAGCACGACAG 1461
Qy 432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
Db 1462 ACTGTTTATCCAGAAATCAGTGTCAATTTGGTTTCATAATGCTAGGTTTCGCTATGG 1521
Qy 452 LysProMetValGluGluMetTyrLeuGluGluValLys----- 464
Db 1522 AAACCAATGTTGGAAGAGATGTATCAACAGAGCAAAAGAAAGAGAGCAGAGAA 1581
Qy 465 -----AenGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLys 480
Db 1582 GAAATGAAATCAACAAACAAAGAGAGCAGCAACAAACAAACAGCAGCAGAA 1641
Qy 481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleThrSerSerLeuLeu 500
Db 1642 CCCAACAC-----AATGAAACAACTTCACGTGCATACCGCACAA----- 1683
Qy 501 GlnAspGlyLeuThrThrGlnAlaGluLeuSerThrSerThrIleSerThrSerPro 520
Db 1684 -----ACTCCA 1689
Qy 521 ThrAlaGlyAlaSerLeuHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
Db 1690 ACGACGATGACATCGACACATCAC----- 1713
Qy 541 AspAsnThrThrThrValAspHisIleGluAsnAlaLysLysGlnArgAsnAsp 560
Db 1714 -----GAAACGAC 1722
Qy 561 MethIleLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580
Db 1723 -----TCTTCATTCCTCTCTCTCCGTC-----GCCGCGCTTCT 1755
Qy 581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600
Db 1756 CACGCGGTTTCAGACGGTTCCCGTCGCCACGTGT----- 1791
Qy 601 GlyArgPheAspProHisAspGlnMetThrAlaAsnPheHis-----Gly 616
Db 1792 -----CAGCAAGACGTGACGTCCACGTCGACGGAGATGT 1830
Qy 617 AsnAsn-----GlyValSerLeuThrLeuGlyLeu 626
Db 1831 GTGAACATCATAGATTGGGACCAACACAGACTGGTGACGTGTCTCTTAGCTTGGCTA 1890
Qy 627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
Db 1891 CGCCACTCTGCAATAT-----CCTGATAGAACACTCTTCTCCGTTAGAGACTTT 1944
Qy 647 Gly 647
Db 1945 GGA 1947

RESULT 13

AD002224
ID AD002224 standard; cDNA; 1983 BP.
XX
AC AD002224;
XX
DT 01-JUL-2004 (first entry)
XX
DE Thalecress transcription factor cDNA #319.
XX
KW Thalecress; transcription factor; ss; gene; plant; transgenic;
KW abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;
KW phosphate limitation; potassium limitation; nitrogen limitation;
KW hormone sensitivity; disease resistance; sugar sensing; seed germination;
KW flowering; inflorescence architectural change;
KW meristem cell differentiation; phyllotaxy; apical dominance;
KW trichome development; seed development; premature senescence;
KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
KW seed morphology; secondary metabolism; light response; shade avoidance.
XX
OS Arabidopsis thaliana.
XX
PN US2004045049-A1.
XX
PD 04-MAR-2004.
XX
PF 10-APR-2003; 2003US-00412699.
XX
PR 13-SEP-1999; 99US-003194519.
PR 21-JAN-2000; 2000US-00489376.
PR 17-FEB-2000; 2000US-00508720.
PR 22-MAR-2000; 2000US-00532591.
PR 22-MAR-2000; 2000US-00533029.
PR 22-MAR-2000; 2000US-00533030.
PR 22-MAR-2000; 2000US-00533392.
PR 22-MAR-2000; 2000US-00533648.
PR 08-APR-2000; 2000WO-US009448.
PR 16-NOV-2000; 2000US-00713994.
PR 27-MAR-2001; 2001US-00819142.
PR 17-APR-2001; 2001US-00837444.
PR 30-JAN-2002; 2002US-00958131.
PR 14-JUN-2002; 2002US-00171468.
PR 09-AUG-2002; 2002US-00225066.
PR 09-AUG-2002; 2002US-00225067.
PR 17-DEC-2002; 2002US-0434166P.
PR 25-FEB-2003; 2003US-00374780.
XX
(ZHAN// ZHANG J.
PA (FROM// FROMM M E.
PA (HEAR// HEARD J E.
PA (RIEC// RIECHMANN J L.
PA (ADAM// ADAM L J.
PA (BROU// BROUN P E.
PA (PINE// PINEDA O.
PA (REUB// REUBER T L.
PA (KEDD// KEDDIE J S.
PA (YUGG// YU G.
PA (JIAN// JIANG C.
PA (SAMA// SAMAHA R S.
PA (PILG// PILGRIM M L.
PA (CREE// CREELMAN R A.
PA (DUBE// DUBELL A N.
PA (RATC// RATCLIFFE O.
PA (KUMI// KUMIMOTO R.
PA (SHER// SHERMAN B K.
XX
PI Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;
PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;
PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
PI Sherman BK;
XX WPI; 2004-225755/21.
DR P-PSDB; AD002225.


```
Db 1282 CTAGACGAGATTGGCTCAGCAACGAGCGCTTTTCATCATATGGGTATGATGGAGCAAGAG 1341
Qy 392 AlaTtArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTtpLeu 411
Db 1342 GCATGGAGACCGACACGCTGTTGCCCTGAACGCTCCGTTAATATCTTAGAGCTTGGCTA 1401
Qy 412 PheGluHisPheLeuHisProTyrProLeuAspSerAspLysIleMetLeuAlaLysGln 431
Db 1402 TTCGAGCATTTCTTAATCCGTACCCCAAGCGATGCTGATAAGCACCTCTTAGCACGACAG 1461
Qy 432 ThrGlyLeuThrArgSerGlnValSerAsnTyrPheIleAsnAlaArgValArgLeuTtp 451
Db 1462 ACTGGTTATCCGAATACAGGTGTCAAATTTGGTTCTAATGCTAGGTTCCGCTATGG 1521
Qy 452 LysProMetValGluGluMetTyrLeuGluValLys----- 464
Db 1522 AAACCAATGTGGAGAGATGTATCAACAGAGCAAAAGAGAGAGAGAGAGAGAA 1581
Qy 465 -----AsnGlnGluInAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
Db 1582 GAAAATGAAATCAACAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1641
Qy 481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
Db 1642 CCCAACAC-----AATGAAACAACTTCACTGTCTATAACCGCACAA----- 1683
Qy 501 GlnAspGlyIleThrThrGlnAlaGluIleSerThrSerThrIleSerThrPro 520
Db 1684 -----ACTCCA 1689
Qy 521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
Db 1690 ACGACGATGATCGACACATCAC----- 1713
Qy 541 AspAsnThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAsp 560
Db 1714 -----GAAAACGAC 1722
Qy 561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580
Db 1723 -----TCTTCATTCTCTCTCCGTC-----GCCGCGCTTCT 1755
Qy 581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600
Db 1756 CACGGCGTTTCAGACGGTTCCCGTCGCCACGCT----- 1791
Qy 601 GlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHis-----Gly 616
Db 1792 -----CAGCAAGACGTCAGTGACATCCACGTCGACGAGATGGT 1830
Qy 617 AsnAsn-----GlyValSerLeuThrLeuGlyLeu 626
Db 1831 GTGACGCTATAGATTCCGGGACCAACACAGCTGGTGACGTGCTCTTACGCTTGGCTA 1890
Qy 627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
Db 1891 CGCCACTCTGCAATATT-----CCTGATAAGAACACTCTTCTCCGTTAGAGACTTT 1944
Qy 647 Gly 647
Db 1945 GGA 1947
RESULT 14
AAC39138
ID AAC39138 standard; DNA; 2025 BP.
XX
AC AAC39138;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23512.
XX
```

```
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
PN
XX
PD
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
```

Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway; metabolic pathway;
promoter; termination sequence; ss.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-00301439.

25-FEB-1999; 99US-0121825P.

05-MAR-1999; 99US-0123180P.

09-MAR-1999; 99US-0123548P.

23-MAR-1999; 99US-0125788P.

25-MAR-1999; 99US-0126264P.

29-MAR-1999; 99US-0126785P.

01-APR-1999; 99US-0127462P.

06-APR-1999; 99US-0128234P.

08-APR-1999; 99US-0128714P.

16-APR-1999; 99US-0129845P.

19-APR-1999; 99US-0130077P.

21-APR-1999; 99US-0130449P.

23-APR-1999; 99US-0130510P.

28-APR-1999; 99US-0130891P.

30-APR-1999; 99US-0131449P.

30-APR-1999; 99US-0132048P.

04-MAY-1999; 99US-0132484P.

05-MAY-1999; 99US-0132485P.

06-MAY-1999; 99US-0132486P.

07-MAY-1999; 99US-0132863P.

11-MAY-1999; 99US-0134256P.

14-MAY-1999; 99US-0134218P.

14-MAY-1999; 99US-0134219P.

14-MAY-1999; 99US-0134221P.

18-MAY-1999; 99US-0134370P.

19-MAY-1999; 99US-0134941P.

20-MAY-1999; 99US-0135124P.

21-MAY-1999; 99US-0135353P.

25-MAY-1999; 99US-0136021P.

27-MAY-1999; 99US-0136392P.

28-MAY-1999; 99US-0136782P.

01-JUN-1999; 99US-0137222P.

03-JUN-1999; 99US-0137528P.

04-JUN-1999; 99US-0137502P.

07-JUN-1999; 99US-0137724P.

08-JUN-1999; 99US-0138094P.

10-JUN-1999; 99US-0138540P.

14-JUN-1999; 99US-0139119P.

16-JUN-1999; 99US-0139452P.

17-JUN-1999; 99US-0139453P.

18-JUN-1999; 99US-0139492P.

18-JUN-1999; 99US-0139454P.

18-JUN-1999; 99US-0139455P.

18-JUN-1999; 99US-0139456P.

18-JUN-1999; 99US-0139457P.

18-JUN-1999; 99US-0139458P.

18-JUN-1999; 99US-0139459P.

18-JUN-1999; 99US-0139460P.

18-JUN-1999; 99US-0139461P.

18-JUN-1999; 99US-0139462P.

18-JUN-1999; 99US-0139463P.

18-JUN-1999; 99US-0139750P.

18-JUN-1999; 99US-0139763P.

21-JUN-1999; 99US-0139817P.

22-JUN-1999; 99US-0139899P.

23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142320P.
PR 12-JUL-1999; 99US-0142377P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.

PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157553P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 3.8e-58 Length: 2025
Score: 820.00 Matches: 222
Percent Similarity: 49.84% Conservative: 92
Best Local Similarity: 35.24% Mismatches: 143
Query Match: 23.01% Indels: 173
DB: 3 Gaps: 23

US-10-624-201A-2 (1-688) x AAC39138 (1-2025)

QY 16 HisGlnGlnArgHisAsn-----HisGlyAsnSerAsnAsnAsn 29
Db 95 CACACTCAACAACAACACTCTTAGGATCTCATTTCCCGACCAACTCATCGAAGTAAGAA 154
QY 30 ileGlnThrLeuTyrLeuMetAsnProAsnAsnTyr-MetGlnGlyTyrThrSerAs 49
Db 155 CTCACCTGTAATAAAGCTTTTAAAGGTGTAAACACCAATGGCAGTTATTATACAAAGTAA 214
QY 49 pThrGln---GlnGlnGlnGlnLeuPheLeu-----AsnSerSerPr 63
Db 215 TGTCGGTTGTACCAGCAAGAACCAATCTTTCTCAACCATCAACAACAAACCAACAGC 274
QY 63 oAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGlnGlnGlnHi 83
Db 275 TTCCTCTTCATCCGCGCGGCTTCTTTCACGTCACCGCGGGGATCTGTTTCGAACAGA 334

PT New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.

PS Claim 1: SEO ID NO 1174: 510pp: English:

The present invention relates to novel plant transcription factor proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The sequences can be used to produce transgenic plants, which overexpress (III), where the transgenic plant has an altered trait as compared to a non-transgenic plant or wild-type plant. The transgenic plant comprises an altered trait selected from increased tolerance to abiotic stress, increased tolerance to osmotic stress, increased tolerance to cold, increased germination in cold, increased tolerance to heat, increased germination in heat, increased tolerance to freezing conditions, increased tolerance to low nitrogen conditions, increased tolerance to low phosphate conditions, increased tolerance to disease, including fungal disease and particularly Fusarium, Botrytis and Botrytis, increased tolerance to multiple fungal pathogens, increased resistance to glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA, increased sensitivity to ACG, altered sugar sensing, increased tolerance to sugars, altered carbon/nitrogen sensing, early flowering, late flowering, altered flower structure, loss of flower determinacy, reduced fertility, altered shoot meristem development, altered branching pattern, altered stem morphology, altered vascular tissue structure, reduced apical dominance, altered trichome density, altered trichome development, altered trichome structure, altered root development, altered shade avoidance, altered seed development, altered seed ripening, altered seed germination, slow growth, fast growth, altered cell differentiation, altered cell proliferation, altered cell expansion, altered phase change, altered senescence, abnormal embryo development, altered programmed cell death, lethality when overexpressed, altered necrosis patterns, increased plant size, increased biomass, large seedlings, dwarfed plants, dark green leaves, change in leaf shape, increased leaf size and mass, light green or gray leaves, glossy leaves, altered abaxial/adaxial polarity, altered seed coloration, altered seed size, altered seed shape, large seed, increased leaf wax, increased leaf fatty acids, altered seed oil content, altered seed protein content, altered seedprenyl content, altered leaf prenyl lipid content, increased anthocyanin levels, and decreased anthocyanin levels. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at <http://wipo.int/pub/published> pct sequences.

Sequence 2458 BP: 751 A: 488 C: 545 G: 674 T: 0 U: 0 Other: XX

Alignment Scores:

Fragment No.:	1.06E-57	Length:	2458
Pred. No.:	Score: 816.00	Matches:	218
	Percent Similarity: 49.9%	Conservative:	83
	Best Local Similarity: 36.1%	Mismatches:	149
	Query Match: 22.9%	Indels:	153
DB:	12	Gaps:	21

US-10-624-201A-2 (1-688) x ADO62707 (1-2458)

Qy	69	LeuCysHisAlaAsnIleGlnHisAlaProLeuGln---GlnGlnHisPheValGlyVal	87
Db	513	CTCATGCCCTTCTAAATGCATCCGTCTCTCCATTTCAATTACCAATATACGACAGGGTGTC	572
Qy	88	-----ProLeuProAlaValSerLeuHisAspGln---	97
Db	573	TTACCACACTAATGAATGATTTTCCAATCCAAAGGGAGCCATGCTCTCAAAGATGATGAG	632
Qy	98	---IleAsnHisHieGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGin	116
Db	633	GCCAAATAACCTACATGGGAGTTTAGA-----	659
Qy	117	ValIleValProSerSerThrClyValSerAlaThrSerCysGlyGlyIleThrThrAsp	136
Db	660	-----AGTGCGTAATGCATGCATCTGTGCTTCTGTGGAGCGTTT---CTTGAC	704
Qy	137	LeuAlaserGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGin	156

```
Db 1620 GTTCAGAGTCTGATTAACATTAAACAATAATAATGGGAAGAGACACAGGACAATTTA 1679
Qy 487 ProAsnGluGluLysHisProIleThrSerLeuLeuGlnAspGlyIleThrThr 506
Db 1680 GTAACCTTCACAGTGTCCAAAGTCCAACTCAAGCAGAACTGGATAGAGGGGTCAATGTA 1739
Qy 507 ThrGlnAlaGluLeuSerThrSerThr----- 515
Db 1740 GAAAAATGTGTGATGGATTCTGGAACCTGGGAAATTGCAGGGTGTATCAAAACCAACAAAGG 1799
Qy 516 -----IleSerThrSerProThr----- 521
Db 1800 CTGGGCATGAACAACAACAACACTTTTATTCTATCTCAACCAACAGAAATGATGGAGGC 1859
Qy 522 ---AlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
Db 1860 CTAATGGTGTGTTCTACTCTATGCCACATATGATTATCAGAGTTGGGTAACTTTACTGTG 1919
Qy 541 AspAsnThrThrThrThrValAspHisIleGluAsn----- 552
Db 1920 GGTAGCCACGTTTCCCTTGCATTGGAGTTGAGGAACCTGTGAAAAGTGTATGGGTTTGCCATG 1979
Qy 553 -----AsnAlaLysLysGlnArgAsnAspMethHisLysPheSerProSerSerIle 569
Db 1980 TCTGATGATGCCATACCGAAACGACGPAATCAAGCATTTGCCCTCTTCCCCGGAGACTGAT 2039
Qy 570 LeuSerSerValAspMetGlu-AlaLysAlaArgGluSerSerAsnLysGlyPheThrAs 589
Db 2040 TTG-----CTAGATTATCATTTACGGATTTCAGGAAGCAACAACACACAGGTTTGGCAAT 2093
Qy 589 nProLeu 591
Db 2094 CCCCAT 2100
```

Search completed: July 29, 2005, 15:05:54
Job time : 1040 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 14:20:37 ; Search time 307 Seconds
(without alignments)
3666.965 Million cell updates/sec

Title: US-10-624-201A-2

Perfect score: 3563

Sequence: 1 MYQGTSDNTNIOADHQHRH.....GNKRFPQLLPDFVTGNLGT 688

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO pool/US10624201/runat 27072005 122958 10931/app query.fasta_1.839
-DB=Issued_Patents_NA -QPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10624201 @CGN 1 1 177 @runat 27072005 122958 10931 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	586.5	16.5	1239	4	US-09-640-211A-65
2	493	13.8	529	4	Sequence 65, Appl
3	445	12.5	381	4	Sequence 77, Appl
4	445	12.5	381	4	Sequence 102, Appl
5	420	11.8	316	4	Sequence 1472, Ap
6	420	11.8	462	4	Sequence 1854, Ap
7	419	11.8	343	4	Sequence 359, Appl
8	411	11.5	323	4	Sequence 1578, Ap
9	360	10.1	698	4	Sequence 418, Appl
10	328.5	9.2	690	4	Sequence 37, Appl
11	319.5	9.0	380	4	Sequence 400, Appl
12	319.5	9.0	380	4	Sequence 111, Appl
					Sequence 1246, Ap

13	297	8.3	366	4	US-09-640-211A-1241	Sequence 1241, Ap
14	260	7.3	171	4	US-09-640-211A-1845	Sequence 1845, Ap
15	256	7.2	260	4	US-09-640-211A-351	Sequence 351, App
16	256	7.2	260	4	US-09-640-211A-1833	Sequence 1833, App
17	251	7.0	474	4	US-09-640-211A-57	Sequence 57, Appl
18	251	7.0	474	4	US-09-640-211A-1399	Sequence 1399, Ap
c	236.5	6.6	486	4	US-09-640-211A-50	Sequence 50, Appl
19	233	6.5	158	4	US-09-640-211A-370	Sequence 370, App
20	226.5	6.4	407	4	US-09-640-211A-383	Sequence 383, App
21	213	6.0	157	4	US-09-640-211A-394	Sequence 394, App
22	208.5	5.9	2905	4	US-09-919-039-381	Sequence 381, App
23	208.5	5.9	2905	4	US-09-976-594-1031	Sequence 1031, Ap
24	208.5	5.9	2905	4	US-09-976-594-1031	Sequence 1031, Ap
25	208	5.8	533	4	US-09-640-211A-185	Sequence 185, App
26	201	5.6	2470	4	US-09-949-016-830	Sequence 830, App
27	198.5	5.6	3190	4	US-09-949-016-5219	Sequence 5219, Ap
28	197.5	5.5	3439	4	US-09-949-016-648	Sequence 648, App
29	197	5.5	3439	4	US-09-976-594-195	Sequence 195, App
30	188	5.3	2200	4	US-09-949-016-3863	Sequence 3863, Ap
31	174	4.9	2049	4	US-09-624-945-2	Sequence 2, Appli
32	172	4.8	273	4	US-09-313-294A-2156	Sequence 2156, Ap
33	171	4.8	1427	4	US-09-949-016-1856	Sequence 1856, Ap
34	168.5	4.7	2952	1	US-08-306-691B-54	Sequence 54, Appl
35	166.5	4.7	6854	1	US-08-468-036-4	Sequence 4, Appli
36	166.5	4.7	6854	2	US-08-376-843-4	Sequence 11081, A
37	166	4.7	1233	4	US-09-270-767-11081	Sequence 5149, Ap
38	165.5	4.6	1086	4	US-09-248-796A-5149	Sequence 4288, Ap
39	163	4.6	1530	4	US-09-949-016-4288	Sequence 460, App
40	163	4.6	1562	4	US-09-949-016-460	Sequence 927, App
41	162.5	4.6	1579	4	US-09-620-312D-927	Sequence 13, Appl
42	161	4.5	2965	4	US-09-665-479A-13	Sequence 94, Appl
43	160	4.5	440	4	US-09-640-211A-94	Sequence 2142, Ap
44	155.5	4.4	282	4	US-09-313-294A-2142	Sequence 4397, Ap
45	154.5	4.3	2537	4	US-09-949-016-4397	

ALIGNMENTS

RESULT 1

US-09-640-211A-65
; Sequence 65, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-65

Alignment Scores:
Pred. No.: 1.7e-56 Length: 1239
Score: 586.50 Matches: 151
Percent Similarity: 50.39% Conservative: 42
Best Local Similarity: 39.43% Mismatches: 97
Query Match: 16.46% Indels: 93
DB: 4 Gaps: 12

US-10-624-201A-2 (1-688) x US-09-640-211A-65 (1-1239)

Qy 20 HisAsnHisGlyAsnSerAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsn 39
Db 240 CACCAGAGCTCATCTGACCCAGATGGAGCCCTACAAACCTCTGCTCTCATGACCCGCC 299

Db 503 CACTCTGTGGATTCAACTACCTG 526

RESULT 3

US-09-640-211A-102

; Sequence 102, Application US/09640211A

; Patent No. 6833446

; GENERAL INFORMATION:

; APPLICANT: Wood, Marion

; APPLICANT: Shenk, Michael A.

; APPLICANT: McGrath, Annette

; APPLICANT: Glenn, Matthew

; TITLE OF INVENTION: Compositions and Methods for the

; TITLE OF INVENTION: Modification of Gene Transcription

; FILE REFERENCE: 11000.1021C1U

; CURRENT APPLICATION NUMBER: US/09/640,211A

; CURRENT FILING DATE: 2000-08-16

; NUMBER OF SEQ ID NOS: 2368

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 102

; SEQ ID NO 102

; LENGTH: 381

; TYPE: DNA

; ORGANISM: Eucalyptus grandis

US-09-640-211A-102

Alignment Scores:

Pred. No.: 2,96e-41 Length: 381

Score: 445.00 Matches: 84

Percent Similarity: 80.80% Conservative: 17

Best Local Similarity: 67.20% Mismatches: 20

Query Match: 12.49% Indels: 4

DB: 4 Gaps: 1

US-10-624-201A-2 (1-688) x US-09-640-211A-102 (1-381)

Qy 344 AlaGluGlnVallyserSerLysSerLysSerLysGluGluGluGlyLeuGlyGly--- 362

Db 3 GCCGCCCACTGAAGCAGACACCGCTCAAGTTCGACGAGCGCGCGGCGCTCC 62

Qy 363 -----LysileGluGlySerArgLeuLysPheValAspHisLeuArgGlnGln 379

Db 63 GGTATAACCAAGCGGCGAGACACCGCTCAAGTTCGACGAGCGCTGAGGCGAG 122

Qy 380 ArgAlaLeuGlnInleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeu 399

Db 123 AGGCGCTTCCACCAGATGGCATGATGGAGCAAGAGCGCTGGAGCGCGCGGCGCTG 182

Qy 400 ProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyr 419

Db 183 CCGAGCGGTCGTCAACATCTGCGTGCATCGCTTCGAGCATTTCTTGATCCCGTAT 242

Qy 420 ProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnVal 439

Db 243 CCAAGTGACGCTGAAGCATCTGTTGGTCGACAGACTGGTCTCTCCAGAAACCGAGTC 302

Qy 440 SerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyr 459

Db 303 TCGAATTGGTTTCATAAATCCAGGGTCCGGTTGTGGAAACCCATGGTGGAGGAGATGTAC 362

Qy 460 LeuGluGluVallys 464

Db 363 CAGCAAGAGTCCAAA 377

RESULT 4

US-09-640-211A-1472

; Sequence 1472, Application US/09640211A

; Patent No. 6833446

; GENERAL INFORMATION:

; APPLICANT: Wood, Marion

; APPLICANT: Shenk, Michael A.

; APPLICANT: McGrath, Annette

; APPLICANT: Glenn, Matthew

; TITLE OF INVENTION: Compositions and Methods for the

; TITLE OF INVENTION: Modification of Gene Transcription

; FILE REFERENCE: 11000.1021C1U

; CURRENT APPLICATION NUMBER: US/09/640,211A

; CURRENT FILING DATE: 2000-08-16

; NUMBER OF SEQ ID NOS: 2368

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1472

; SEQ ID NO 1472

; LENGTH: 381

; TYPE: DNA

; ORGANISM: Eucalyptus grandis

US-09-640-211A-1472

Alignment Scores:

Pred. No.: 1.52e-38 Length: 316

Score: 445.00 Matches: 84

Percent Similarity: 80.80% Conservative: 17

Best Local Similarity: 67.20% Mismatches: 20

Query Match: 12.49% Indels: 4

DB: 4 Gaps: 1

US-10-624-201A-2 (1-688) x US-09-640-211A-1472 (1-381)

Qy 344 AlaGluGlnVallyserSerLysSerLysSerLysGluGluGluGlyLeuGlyGly--- 362

Db 3 GCCGCCCACTGAAGCAGACACCGCTCAAGTTCGACGAGCGCGCGGCGCTCC 62

Qy 363 -----LysileGluGlySerArgLeuLysPheValAspHisLeuArgGlnGln 379

Db 63 GGTATAACCAAGCGGCGAGACACCGCTCAAGTTCGACGAGCGCTGAGGCGAG 122

Qy 380 ArgAlaLeuGlnInleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeu 399

Db 123 AGGCGCTTCCACCAGATGGCATGATGGAGCAAGAGCGCTGGAGCGCGCGGCGCTG 182

Qy 400 ProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyr 419

Db 183 CCGAGCGGTCGTCAACATCTGCGTGCATCGCTTCGAGCATTTCTTGATCCCGTAT 242

Qy 420 ProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnVal 439

Db 243 CCAAGTGACGCTGAAGCATCTGTTGGTCGACAGACTGGTCTCTCCAGAAACCGAGTC 302

Qy 440 SerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyr 459

Db 303 TCGAATTGGTTTCATAAATCCAGGGTCCGGTTGTGGAAACCCATGGTGGAGGAGATGTAC 362

Qy 460 LeuGluGluVallys 464

Db 363 CAGCAAGAGTCCAAA 377

RESULT 5

US-09-640-211A-1854

; Sequence 1854, Application US/09640211A

; Patent No. 6833446

; GENERAL INFORMATION:

; APPLICANT: Wood, Marion

; APPLICANT: Shenk, Michael A.

; APPLICANT: McGrath, Annette

; APPLICANT: Glenn, Matthew

; TITLE OF INVENTION: Compositions and Methods for the

; TITLE OF INVENTION: Modification of Gene Transcription

; FILE REFERENCE: 11000.1021C1U

; CURRENT APPLICATION NUMBER: US/09/640,211A

; CURRENT FILING DATE: 2000-08-16

; NUMBER OF SEQ ID NOS: 2368

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1854

; SEQ ID NO 1854

; LENGTH: 316

; TYPE: DNA

; ORGANISM: Pinus radiata

US-09-640-211A-1854

Alignment Scores:

Pred. No.: 1.52e-38 Length: 316

Score: 445.00 Matches: 84

Percent Similarity: 80.80% Conservative: 17

Best Local Similarity: 67.20% Mismatches: 20

Query Match: 12.49% Indels: 4

DB: 4 Gaps: 1

Score: 420.00 Matches: 79
Percent Similarity: 87.38% Conservative: 11
Best Local Similarity: 76.70% Mismatches: 11
Query Match: 11.79% Indels: 2
DB: 4 Gaps: 1

US-10-624-201A-2 (1-688) x US-09-640-211A-1854 (1-316)

QY 380 ArgAlaLeuGlnGlnIleGlyMetGlnProAsnAlaTrpArgProGlnArgGlyLeu 399
Db 2 CGGGCTCTCCAAATATAGCATGATTCACGAGCATGCTTGGAGGCCACAGAGAGACTT 61
QY 400 ProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyr 419
Db 62 CCGAGCGATCTGTTCTGCTTACGGCTTGGCTATTTGAACATTTCTTCATCCGTAT 121
QY 420 ProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnVal 439
Db 122 CCAAAAGATCCAGACAAACATATGCTCGCGAGACAGACTGGGCTTACCAGAAATCAGGTC 181
QY 440 SerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyr 459
Db 182 TCAAAATGGTTTATAAATGACGTGACGCTCTGGAAGCCTATGTTGGAAGAGATGTAT 241
QY 460 LeuGluGluValLysAsnGlnGlnAsnSerThrAsnThrSerGlyAspAsnLysAsn 479
Db 242 GTGGAGAAACAAAGGAGGCGAGAGTAGAC-----CATGGATCAAAATGATAAAACAGGT 295
QY 480 LysGluThr 482
Db 296 AAGGAGAGT 304

RESULT 6

US-09-640-211A-359
; Sequence 359, Application US/09640211A
; Patent No. 6833446

; GENERAL INFORMATION:
; APPLICANT: Wood, Marion

; APPLICANT: Shenk, Michael A.

; APPLICANT: McGrath, Annette

; APPLICANT: Glenn, Matthew

; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription

; FILE REFERENCE: 11000.1021C1U

; CURRENT APPLICATION NUMBER: US/09/640,211A

; CURRENT FILING DATE: 2000-08-16

; NUMBER OF SEQ ID NOS: 2368

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 359

; TYPE: DNA

; ORGANISM: Pinus radiata

US-09-640-211A-359

Alignment Scores:
Pred. No.: 2,99e-38 Length: 462
Score: 420.00 Matches: 79
Percent Similarity: 87.38% Conservative: 11
Best Local Similarity: 76.70% Mismatches: 11
Query Match: 11.79% Indels: 2
DB: 4 Gaps: 1

US-10-624-201A-2 (1-688) x US-09-640-211A-359 (1-462)

QY 380 ArgAlaLeuGlnGlnIleGlyMetGlnProAsnAlaTrpArgProGlnArgGlyLeu 399
Db 2 CGGGCTCTCCAAATATAGCATGATTCACGAGCATGCTTGGAGGCCACAGAGAGACTT 61
QY 400 ProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyr 419
Db 62 CCGAGCGATCTGTTCTGCTTACGGCTTGGCTATTTGAACATTTCTTCATCCGTAT 121
QY 420 ProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnVal 439

Db 122 CCAAAAGATCCAGACAAACATATGCTCGCGACAGACACTGGGCTTACCAGAAATCAGGTC 181
QY 440 SerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyr 459
Db 182 TCAAAATGGTTTATAAATGACGTGACGCTCTGGAAGCCTATGTTGGAAGAGATGTAT 241
QY 460 LeuGluGluValLysAsnGlnGlnAsnSerThrAsnThrSerGlyAspAsnLysAsn 479
Db 242 GTGGAGAAACAAAGGAGGCGAGAGTAGAC-----CATGGATCAAAATGATAAAACAGGT 295
QY 480 LysGluThr 482
Db 296 AAGGAGAGT 304

RESULT 7

US-09-640-211A-1578
; Sequence 1578, Application US/09640211A
; Patent No. 6833446

; GENERAL INFORMATION:

; APPLICANT: Wood, Marion

; APPLICANT: Shenk, Michael A.

; APPLICANT: McGrath, Annette

; APPLICANT: Glenn, Matthew

; TITLE OF INVENTION: Compositions and Methods for the

; TITLE OF INVENTION: Modification of Gene Transcription

; FILE REFERENCE: 11000.1021C1U

; CURRENT APPLICATION NUMBER: US/09/640,211A

; CURRENT FILING DATE: 2000-08-16

; NUMBER OF SEQ ID NOS: 2368

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1578

; LENGTH: 343

; TYPE: DNA

; ORGANISM: Pinus radiata

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(343)

; OTHER INFORMATION: n = A,T,C or G

US-09-640-211A-1578

Alignment Scores:
Pred. No.: 2,28e-38 Length: 343
Score: 419.00 Matches: 77
Percent Similarity: 82.46% Conservative: 17
Best Local Similarity: 67.54% Mismatches: 18
Query Match: 11.78% Indels: 2
DB: 4 Gaps: 1

US-10-624-201A-2 (1-688) x US-09-640-211A-1578 (1-343)

QY 365 GluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGln 384
Db 1 GAAACACCAAGGTTGGGATNTCTAGAACAAGACATACGACAAACAGCGCCATTTCCACCAC 60
QY 385 IleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaVal 404
Db 61 TTAGGATGTATGAGCAGCAGCACCCCTTGGCGACCGCAGAGAGGACTTCTCTGAAACGCTCTG 120
QY 405 SerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAsp 424
Db 121 TCTGTTCTTCGTGCATGTTGTTTGAGCATTTCTTCACCCCGCTATCCAACTGATGCAGAT 180
QY 425 LysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIle 444
Db 181 AAGCATATATTGGCTAAGCAAACTGGCTTACAAGAAAGTCAGGTATCAAAATGGTTTATA 240
QY 445 AsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluValLys 464
Db 241 AATGCCAGGTTAGACTATGGAAGCCCATGGTGGAGGAGATGTACATGGAAGAACTCAAG 300
QY 465 AsnGln-----GluGlnAsnSerThrAsnThrSerGlyAsp 476
::: ||| ::| |||::| ::|


```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 698
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-37

Alignment Scores:
Pred. No.: 4,41e-31      Length: 698
Score: 360.00           Matches: 74
Percent Similarity: 69.23%      Conservative: 34
Best Local Similarity: 47.44%    Mismatches: 46
Query Match: 10.10%            Indels: 2
DB: 4                     Gaps: 2

US-10-624-201A-2 (1-688) x US-09-640-211A-37 (1-698)
QY 203 MetValLeuGlySerLysTyLeuLysAlaGlnGluLeuLeuAspGluValValAsn 222
Db 222 GTCATTCCTCAATTCCTAGGTACTTGGGGTGGCGGACCACTGCTTGTATGAATAGTGAAT 281
QY 223 IleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsnLysGlu 242
Db 282 ---GTGCAAGAGGCTTTGAAGCGCTCTGATGACGGAATGACCAATCTAGACAT---GAG 335
QY 243 SerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSerSer 262
Db 336 AACCAAGAGAGCCCAAGATGCGGATGGGGGTTCACGAAGACGAAGCATCTCTCAATCCC 395
QY 263 ArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMet 282
Db 396 CAAGAATCGGCAGTAACTCTAGCGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 455
QY 283 LysLysAlaLysLeuLeuAlaMetLeuGluValGluGlnArgGlyArgGlnTyHis 302
Db 456 AAGCTCAGCAAGCTCTTGTCCATGTTGACGAGGTTGATAAAGGTACAAAGCAGTACTAT 515
QY 303 HisGlnMetGlnIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLys 322
Db 516 CACCAATGCGAGATCGTGTACAGTCTTTTATACATAGCAGGAAGCGGTGCGACGCAAG 575
QY 323 SerTyThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAla 342
Db 576 CCCTACACGGCGTTGGCTCCAGAGGATATCCGCCACATTCGCGTCCCTGCATGACGCG 635
QY 343 IleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGlu 358
Db 636 ATCAGCGGTCAAAATTCAGCAACCCGTAAGAAAGTCTCGGAGAGCAAGAC 683

RESULT 10
US-09-640-211A-400
; Sequence 400, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 400
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-400

Alignment Scores:
Pred. No.: 1.7e-27      Length: 690

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 698
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-37

Alignment Scores:
Pred. No.: 4,41e-31      Length: 698
Score: 360.00           Matches: 74
Percent Similarity: 69.23% Conservative: 34
Best Local Similarity: 47.44% Mismatches: 46
Query Match: 10.10%      Indels: 2
DB: 4                   Gaps: 2

US-10-624-201A-2 (1-688) x US-09-640-211A-37 (1-698)
QY 203 MetValLeuGlySerLysTyLeuLysAlaGlnGluLeuLeuAspGluValValAsn 222
Db 222 GTCATTCCCAATTCCTAGGTACTTGGGGTGGCGGACCACTGCTTGTATGAATAGTGAAT 281
QY 223 lleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsnLysGlu 242
Db 282 ---GTGCAAGAGGCTTTGAAGCGCTCTGATGACGGAATGACCAATCTAGACAT---GAG 335
QY 243 SerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSerSer 262
Db 336 AACCAAGAGAGCCCAAGATGCGGATGGGGGTTCACGAAGACGAAGCATCTCTCAATCCC 395
QY 263 ArgGlnLysAsnGluValAlaValGluLeuThrAlaGlnArgGlnGluLeuGlnMet 282
Db 396 CAAGAATCGGCAGTAACTCTAGCGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 455
QY 283 LysLysAlaLysLeuLeuAlaMetLeuGluValGluGlnArgGlyArgGlnTyHis 302
Db 456 AAGCTCAGCAAGCTCTTGTCCATGTTGACGAGGTTGATAAAGGTACAAAGCAGTACTAT 515
QY 303 HisGlnMetGlnIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLys 322
Db 516 CACCAATGCGAGTCGTGGTACAGTCTTTTGTATACATAGCAGGAACGGGTGCGACCAAG 575
QY 323 SerTyThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAla 342
Db 576 CCCTACACGGCGTTGGCTCCAGAGGATATCCGCCACATTCGCGTCCCTGCATGACGCG 635
QY 343 lleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGlu 358
Db 636 ATCAGCGGTCAATTCAGCAACCCGTAAGAAAGTCTCGGAGGACGAC 683

RESULT 10
US-09-640-211A-400
; Sequence 400, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 400
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-400

Alignment Scores:
Pred. No.: 1.7e-27      Length: 690

```

Score: 328.50 Matches: 79
Percent Similarity: 61.64% Conservative: 19
Best Local Similarity: 49.69% Mismatches: 46
Query Match: 9.22% Indels: 15
DB: 4 Gaps: 5

US-10-624-201A-2 (1-688) x US-09-640-211A-400 (1-690)

```
QY 181 SerSerSerSerProArgThrAsnValThrIleArgGlyThr---LeuAspGlySer 199
DB 214 AGTGTGAATGGCCCTGAGTGTAACAAACCGATGCATCATGTACGAATGCAATGGAGCA 273
QY 200 Ser-----SerAsnMetValLeuGlySerLeuLeuLeuLeuLeuLeuLeuLeu 217
DB 274 GCAGAGCTCTCAACACACTTAACTGGTTCCAAATATTTAAAGCAGCAGCAATTAATT 333
QY 218 AspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsn 237
DB 334 GATGAAGTTGTAAT---GTTGGAAGGGTATCAAGTCTGATTCACTCAACCATCAAAA 390
QY 238 SerMetAsn-----LysGluSerMetProLeuAlaSerAspValAsnThr 252
DB 391 TCCAAACATGTTGGTGGCAATATCTGACAAAAGAAATATTGCAACTGAAGCTACTACA 450
QY 253 AsnSer-----SerGlyGlyGlyGluSerSerSerArgGlnLysAsn 266
DB 451 AATGACCGAACACATCTGCAATAACAGAGCTTCAATTTCTGCAGAAATGTAATAAAC 510
QY 267 GluValAlaValGluLeuThrAlaGlnArgGlnGlnGluLeuGlnMetLysAlaLys 286
DB 511 GAGCATGCTTTTGGACTCACACAGCTGATAGACAGAACTTCAGATGAAAAAGGCAGAG 570
QY 287 LeuLeuAlaMetLeuGluValGluGlnArgTyrArgGlnTyrHisGlnMetGln 306
DB 571 CTGTGTGCCATGTGTGATGAGGTGATCGAAGGTACAGACAGTACTATCATCAGATGCAA 630
QY 307 IleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThr 325
DB 631 ATCGTTGTTTCATCGTTTGAGACCGCAGCTGGATTTGGGGCTGCCAAGACATACACT 687
```

RESULT 11

US-09-640-211A-111
; Sequence 111, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-111

Alignment Scores:
Pred. No.: 6,27e-27 Length: 380
Score: 319.50 Matches: 69
Percent Similarity: 68.60% Conservative: 14
Best Local Similarity: 57.02% Mismatches: 23
Query Match: 8.97% Indels: 15
DB: 4 Gaps: 4

US-10-624-201A-2 (1-688) x US-09-640-211A-111 (1-380)

```
QY 411 LeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLys 430
```

```
DB 2 CTCCTCGAACACTTTCACCCCTTACCGAAGGATTCGGACAAAGTCATGCTGGCCAAA 61
QY 431 GlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeu 450
DB 62 CAGACAGGGCTCACTAGAACCCAGGTCGAAATGGTTTATAAATGCTCGAGTTCGGCTT 121
QY 451 TrpLysProMetValGluGluMetTyrLeuGluValLysAsnGlnGluGlnAsn--- 469
DB 122 TCGAAGCCCATGTGGAGAGATGTACACGGAGGAAATCAAGAGCAAGAACAGAAATGGG 181
QY 470 -----SerThrAsnThrSerGlyAspAsnLysGluThr 482
DB 182 GGAGGAGCAGAGGAAAAACCAAGCAAGAGTGAACCGGAGGACTCAGCATCCCAAGTCCTCT 241
QY 483 AsnIle-----SerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
DB 242 GGCCTCCAGGACAGGCCCCCAACTCCAATGAGAAACAGC-----ACCAAGAGCTTCAA 295
QY 501 GlnAspGlyIleThrThrThrGlnAlaGlu-----IleSerThrSerThrIleSer 517
DB 296 CCAAGAGGATCACCTCGAGGAACACGACCCCTGCCATCTCTACTAATTCGGCTTCC 355
QY 518 Thr 518
DB 356 TCC 358
```

RESULT 12

US-09-640-211A-1246
; Sequence 1246, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1246
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1246

Alignment Scores:
Pred. No.: 6,27e-27 Length: 380
Score: 319.50 Matches: 69
Percent Similarity: 68.60% Conservative: 14
Best Local Similarity: 57.02% Mismatches: 23
Query Match: 8.97% Indels: 15
DB: 4 Gaps: 4

US-10-624-201A-2 (1-688) x US-09-640-211A-1246 (1-380)

```
QY 411 LeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLys 430
DB 2 CTCCTCGAACACTTTCACCCCTTACCGAAGGATTCGGACAAAGTCATGCTGGCCAAA 61
QY 431 GlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeu 450
DB 62 CAGACAGGGCTCAC TAGAAGCCAGGTCGAAATGGTTTATAAATGCTCGAGTTCGGCTT 121
QY 451 TrpLysProMetValGluGluMetTyrLeuGluValLysAsnGlnGluGlnAsn--- 469
DB 122 TCGAAGCCCATGTGGAGGAGATGTACACGGAGGAAATCAAGAGCAAGAACAGAAATGGG 181
QY 470 -----SerThrAsnThrSerGlyAspAsnLysGluThr 482
```

```
Db 182 GGAGGACGACGAGGAAACCAAGACAGAGTGACAGCGGAGCTCAGCATCCAAGTCTCT 241
Qy 483 AsnIle-----SerAlaProAsnGluGluLysHisProIleThrSerSerLeuLeu 500
Db 242 GGCTCCAGGACAGGCCCCCACTCCAATGAGAACAGC-----ACCAAGAGCTTCAAA 295
Qy 501 GlnAspGlyIleThrThrGlnAlaGlu-----IleSerThrSerThrIleSer 517
Db 296 CCAAGGAGATCACTCGAGGAACACGACACCCCTGCCATCTCTACTAATCGGCTTCC 355
Qy 518 Thr 518
Db 356 TCC 358

RESULT 13
US-09-640-211A-1241
; Sequence 1241, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1241
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1241

Alignment Scores:
Pred. No.: 2,17e-24 Length: 366
Score: 297.00 Matches: 65
Percent Similarity: 64.62% Conservative: 19
Best Local Similarity: 50.00% Mismatches: 30
Query Match: 8.34% Indels: 16
DB: 4 Gaps: 4

US-10-624-201A-2 (1-688) x US-09-640-211A-1241 (1-366)

Qy 413 GluHisPheLeuHisProTyProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 1 GAGCATTTTCTTCATCTTACCCAAAGAGATTCTGACAAAATCTCTTGCAGGACGACA 60
Qy 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db 61 GGCTTGACAAGAGTCAGGCTCGAATTGGTTTCATCAATGCAAGAGTGGCTCTCTGGAAA 120
Qy 453 ProMetValGluGluMetTyLeuGluValLysAsnGlnGluGlnAsnSerThrAsn 472
Db 121 CCTATGTCGAAGAAATGTACAAAGAGAGATTGGGATCGGGAATGGACTCCAACTCA 180
Qy 473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
Db 181 TCTCCGACACAGCCAAAGCAAAACAGGAGATATCAAGTCTCCATGGAGGACCGG--- 237
Qy 493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGln----- 508
Db 238 -----GTGGAAGAAGTGCACACAGAGTTCAACAGCTACACAGATGCAGCTCA 285
Qy 509 AlaGluLeSerThrSerThrIleSerThrSerPro-----ThrAlaGlyAlaSer 525
Db 286 GGCCAGCTCATGGACTCATCTACCGGACTCCAGATGTGCAAAATGGCAGGC----- 339
Qy 526 LeuHisIleAlaHisAsnPheSerPheLeu 535
Db 340 -----CACTCTGTGGGATTCAACTACCTG 363
```

```
RESULT 14
US-09-640-211A-1845
; Sequence 1845, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1845
; LENGTH: 171
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1845

Alignment Scores:
Pred. No.: 9.38e-21 Length: 171
Score: 260.00 Matches: 47
Percent Similarity: 92.86% Conservative: 5
Best Local Similarity: 83.93% Mismatches: 4
Query Match: 7.30% Indels: 0
DB: 4 Gaps: 0

US-10-624-201A-2 (1-688) x US-09-640-211A-1845 (1-171)

Qy 367 SerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnIleGly 386
Db 4 TCCGCTTCTCACTTGTGATCAACAAATTACGACACAGCGAGCTCTTCAGCAGCTAGGA 63
Qy 387 MetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerVal 406
Db 64 ATGATACAGCAGCATGCTGGAGACCACCAAGAGGGCTTCAGAGAGGGCGGTTCTATT 123
Qy 407 LeuArgAlaTrpLeuPheGluHisPheLeuHisProTyProLysAsp 422
Db 124 CTCGGGGCTTGGCTATTGAGCAATTCCTTCATCCGTACCCCAAAAT 171

RESULT 15
US-09-640-211A-351
; Sequence 351, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 351
; LENGTH: 260
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-351

Alignment Scores:
Pred. No.: 5.66e-20 Length: 260
Score: 256.00 Matches: 54
Percent Similarity: 79.49% Conservative: 8
Best Local Similarity: 69.23% Mismatches: 7
Query Match: 7.18% Indels: 9
```

```

DB:          4          Gaps:          3
US-10-624-201A-2 (1-688) x US-09-640-211A-351 (1-260)
QY 413 GluHisPheLeu-HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnTh 432
Db 9 CAGCAATTCCTTCATCCGCTACCCAAAGATGCTGACAAACATTCTAGCAAGACAGAC 68
QY 432 rGlyLeuThrArgSerClnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLy 452
Db 69 TGGACTGACCAAGCCAGGTTTCAAAATTGGTTTATAAATGCACAGTGTCCGCTTTGGAA 128
QY 452 sProMetValGluGluMetTyrLeuGluValLysAsnGlnGlu---GlnAsnSer-- 470
Db 129 ACCCATGGTGGAGAAATGTATATGGAGGAACCTAGAGAGCCGCAACACAGAAATCATGC 188
QY 471 -----ThrAsnThrSerGlyAspAsnLysAsnLysGluThr 482
Db 189 AGCAGATTCCGAGGTAACAAACAGAAAGTGTCAAAAC---AATGAAGAAACG 237

```

Search completed: July 29, 2005, 18:59:21
Job time : 316 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 14:48:42 ; Search time 1102 Seconds
(without alignments)
4040.818 Million cell updates/sec

Title: US-10-624-201a-2

Perfect score: 3563

Sequence: 1 MYQGSTDNITQADHQRRH.....GNKRFFPTQLLPDFVTGNLGT 688

Scoring table:

BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10624201/runat_27072005_122959_10961/app_query.fasta_1.839
-DB=Published Applications_NA_QFT=fastap -SURFIX=rnnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10624201 @CGN 1 1 789 @runat_27072005_122959_10961
-NCPU=6 -ICPU=3 -NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DEUXT=7

Database :

Published Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10A_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1353	38.0	3033	18	US-10-424-599-130628 Sequence 130628, A
2	1174.5	33.0	1555	18	US-10-425-114-13353 Sequence 13353, A
3	1150.5	32.3	2043	9	US-09-938-842A-1337 Sequence 1337, Ap
4	1150.5	32.3	2043	11	US-09-938-842A-1337 Sequence 1337, Ap
5	1150.5	32.3	2043	21	US-10-495-918-153 Sequence 153, App
6	1148.5	32.2	2385	17	US-10-225-066A-335 Sequence 335, App
7	1148.5	32.2	2385	17	US-10-374-780A-2671 Sequence 2671, App
8	1148.5	32.2	2385	22	US-10-225-066A-335 Sequence 335, App
9	1079.5	30.3	1822	18	US-10-424-599-133558 Sequence 133558, A
10	1078	30.3	1606	18	US-10-425-114-7804 Sequence 7804, Ap
11	1050.5	29.5	3029	19	US-10-437-963-60154 Sequence 60154, A
12	1037	29.1	1892	18	US-10-424-599-63874 Sequence 63874, A
13	999	28.0	2313	18	US-10-425-114-15195 Sequence 15195, A
14	999	28.0	2970	20	US-10-425-115-97705 Sequence 97705, A
15	984	27.6	1716	18	US-10-425-114-9882 Sequence 9882, Ap
16	984	27.6	3311	18	US-10-424-599-141176 Sequence 141176, A
17	945	26.5	3028	19	US-10-437-963-41007 Sequence 41007, A
18	941.5	26.4	1498	18	US-10-425-114-33974 Sequence 33974, A
19	926.5	26.0	2799	19	US-10-437-963-31585 Sequence 31585, A
20	912.5	25.6	2943	20	US-10-425-115-118586 Sequence 118586, A
21	864	24.2	1649	18	US-10-425-114-20970 Sequence 20970, A
22	845.5	23.7	3135	19	US-10-437-963-67929 Sequence 67929, A
23	824	23.1	1884	9	US-09-938-842A-905 Sequence 905, App
24	824	23.1	1884	11	US-09-938-842A-905 Sequence 905, App
25	824	23.1	1983	14	US-10-286-264-17 Sequence 17, Appl
26	824	23.1	1983	17	US-10-225-066A-819 Sequence 819, App
27	824	23.1	1983	17	US-10-374-780A-2481 Sequence 2481, Ap
28	824	23.1	1983	22	US-10-412-699B-637 Sequence 637, App
29	824	23.1	1983	22	US-10-225-066A-819 Sequence 819, App
30	823.5	23.0	1905	19	US-10-437-963-89736 Sequence 89736, A
31	818.5	23.0	1645	18	US-10-424-599-65908 Sequence 65908, A
32	810.5	22.7	1637	18	US-10-425-114-13359 Sequence 13359, A
33	802	22.5	1599	9	US-09-938-842A-2154 Sequence 2154, Ap
34	802	22.5	1599	11	US-09-938-842A-2154 Sequence 2154, Ap
35	797.5	22.4	1727	18	US-10-425-114-34958 Sequence 34958, A
36	797.5	22.4	2545	17	US-10-225-068-235 Sequence 235, App
37	797.5	22.4	2545	17	US-10-374-780A-2329 Sequence 2329, Ap
38	797.5	22.4	2545	18	US-10-412-699B-267 Sequence 267, App
39	797.5	22.4	2545	21	US-10-225-068-235 Sequence 235, App
40	795	22.3	2486	18	US-10-424-599-55280 Sequence 55280, A
41	793.5	22.3	1736	20	US-10-425-115-55285 Sequence 55285, A
42	784	22.0	2736	18	US-10-424-599-37574 Sequence 37574, A
43	776.5	21.8	831	19	US-10-767-701-10301 Sequence 10301, A
44	775.5	21.8	1005	18	US-10-425-114-22835 Sequence 22835, A
45	775.5	21.8	2442	18	US-10-424-599-139178 Sequence 139178, A

ALIGNMENTS

RESULT 1

US-10-424-599-130628
; Sequence 130628, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 130628

```
/ LENGTH: 3033
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_88966C.1
US-10-424-599-130628

Alignment Scores:
Pred. No.: 2,36e-124 Length: 3033
Score: 1533.00 Matches: 350
Percent Similarity: 56.87% Conservative: 93
Best Local Similarity: 44.93% Mismatches: 160
Query Match: 37.97% Indels: 176
DB: 18 Gaps: 36

US-10-624-201A-2 (1-688) x US-10-424-599-130628 (1-3033)
Qy 16 HisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsn----- 29
Db 245 CACTGAGAGAGATACAAATCATAGAAATTGAAGAAGAACGCAACATCGCGAGCTACTTTAAT 304
Qy 30 -----IleGlnThrLeuTyrLeuMet 36
Db 305 GGAAGCACTTCGGAATTCACACTTCTGCAGATGGCGGCTTGCAAAACGCTTTACCTTATG 364
Qy 37 AsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnLeu 56
Db 365 AACCCA-----ACTTACGTA-----CCCTAGCTGACGCGCGCCACCACCAACACG 412
Qy 57 LeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHis 76
Db 413 CTCTCGTCAAT-----CCCAACGCACCAACCTC-----GCGAATATCCACAC 457
Qy 77 AlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAsp 96
Db 458 GCG-----CCGCCAGTGTCAACCAACCA 481
Qy 97 GlnIleAsnHis-----HisGlyLeuLeuGlnArgMet-----TrpAsnAsnGlnAsp 112
Db 482 CAACATCACCACGCTGATTACGCGGTGCACAGCATCATCGGATCCGGCAATTCGGATGAC 541
Qy 113 GlnSerGlnGlnValIleVal-----ProSerSer 122
Db 542 CACTCCCGACCGCTGTGATCGCGGAAAAACATCGCGCCTTCACAGCTTCTCCGCGGG 601
Qy 123 ThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGlnLeuAla 142
Db 602 GCAGGCACCTGCCTCCACCAACCGCCCTTCCTACTATAACCTCTGGGGTCCGGTGGTT 661
Qy 143 PheGlnArgProIleProThrProGln-----HisArgGlnGlnGln 157
Db 662 GATCAGCCGGGAACGGCTGTACAGCTCCGGTGGAGGGCGGTTTCGGCGCGCGAGTCAG 721
Qy 158 GlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsn 177
Db 722 CAG-----GGTTGTCTCTCAGCTTTTCATCAG----- 751
Qy 178 AsnAsnIleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAsp 197
Db 752 -----CAACACAAATTTTAGATCAGTTTCCCGCGAGCTTGAC 787
Qy 198 -----GlySerSer----- 200
Db 788 GTGGCCGGACAGCGCCACGTTGCGGNAATGGAACTCGCCGATGTCCGGCTCGATCGGA 847
Qy 201 ---SerAsnMetValLeuGlySerIleTyrLeuLeuAlaAlaGlnGlnLeuLeuAspGlu 219
Db 848 GTTTCAGGTGTTATTATGGGGTCAAAAGTACTTGAAGCTGCACAGAGCTTCTGGATGAA 907
Qy 220 ValValAsnIleValGlyIleSerIleLeuGlyAspAspGlnIleIleAspAsn---Ser 238
Db 908 GTTGTGAAC---GTGGGAAAGGAATCTTCAAAAGAGGAGAAATTTCTCAGAGAGCGTGAA 964
```

```
Qy 239 MetAsnLysGluSer-----MetProLeuAlaSerAspValAsnThrAsnSerSerGly 256
Db 965 GCGAATAGGAATCACTAAATTCTGGTGTCTGGTGATGGTGAGATGGTTCAAGTGT 1024
Qy 257 GlyGlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGln 276
Db 1025 CGAGGAGAAACAGTGCAGGAAACAA-----GTGGTTGAACCTCAGCACCGCTCAG 1075
Qy 277 ArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluValGln 296
Db 1076 AGGCAAGAGCTTCAGATGAAGAAGTCCAAGCTTGACCATCTCGATCAGGTAGAACAA 1135
Qy 297 ArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAla 316
Db 1136 AGGTACAGACATATCACCCCAATGCAATTTGGTATCGTCATTTGAGCAAGCAGCA 1195
Qy 317 GlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPhe 336
Db 1196 GGTATATGGGGCGCAAAATCTTACACTGCGCTTGCCCTAAAAAACTATCTCAAGCAATTC 1255
Qy 337 ArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu 356
Db 1256 AGGTGCTTGAAGATGCAATCTCTGCACAAATCAAAAGCCACGACGAAGACATTGGGTGAA 1315
Qy 357 GluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisLeu 376
Db 1316 GATGATTGTTAGGAGTTAAAGTAGAAGGCTCAAGGCTTAGGTATGTTGACCACCATCTC 1375
Qy 377 ArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGln 396
Db 1376 AGGCAACAAAGGCACCTTCAGCAGCTTAGGAATGATTAACCAATCATGAGGCGCCCAA 1435
Qy 397 ArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeu 416
Db 1436 AGAGGCTTGCCCTGAACAGAGCTTTTCCATCTTCGGGCTTGGCTTTTTCAGCATTTCCCT 1495
Qy 417 HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArg 436
Db 1496 CACCTTATCCAAAGGACTCCGATAAGTTATGCTTAAACAACTCGACTTGTAGG 1555
Qy 437 SerGlnValSerAsnTrpPheIleAlaArgValArgLeuTrpLysProMetValGlu 456
Db 1556 AGCCAGGTGTCAAACTGGTTTATCAATGCCGAGTTCGGCTTGGGAAGCAATGGTTGAA 1615
Qy 457 GluMetTyrLeuGluGluValLysAsnGlnGluGln-----AsnSerThrAsnThrSer 474
Db 1616 GAAATGTACTTGGAGAATAATCAAGGAGCATGAACAGGTAATGGCTCAAGAAACACCAA 1675
Qy 475 GlyAspAsnLysAsnLysGlu-----ThrAsnIleSerAlaProAsn---Glu 489
Db 1676 TCAAAAGAAATCGAGCAAGGAGTTGGCTTCAACAGCAAAATGTTGCATCATCTCCAA 1735
Qy 490 GluLysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAla 509
Db 1736 TCAAAACATGAAGCTTCAACCAACCAA-----AATACTTCCCAACC 1777
Qy 510 GluIleSerThrSerThrIleSerThrProThrAlaGlyValSerLeuHisAla 529
Db 1778 GAGATCTCAACAGCTCT---TCAATGTCTCCAAATGGGAGGTCCTTCAATCCCACTCT 1834
Qy 530 HisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHis 549
Db 1835 ---GGTTTCCATCTTCAGGATCATCTGACATG-----CAA 1867
Qy 550 IleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIle 569
Db 1868 ATTAGAGGTCCAAACCAACCAAGAGTCTGAAATGCAAAAC---TCCCAAGTAGTATC 1924
Qy 570 LeuSerSerValAspMetGluAlaLys-----AlaArgGluSer 582
Db 1925 CTT---TCAGTGCACATGGAATGAAGCATAGTGTGACCATGGAATATAGAGATGCC 1981
Qy 583 SerAsnLys-----GlyPheThrAsnProLeu 591
```

```

Db 1982 AACACAAAGTTTGGTATTGAAAGGCATCACCACAAAGATGGGTAC-----CCTTTG 2035
Qy 592 Met-----AlaAlaTyrAlaMetGlyAspPheGly 601
Db 2036 ATGACTAGCAATCCAAACCATGGTGGGGGTTTGGAGCATTCACCATGGAGACATTGGA 2095
Qy 602 ---ArgPheAspProHisAspGlnMetThrAlaAsnPheHisGlyAsnAenglyVal 620
Db 2096 AGCAGATTAATGCTACTACTAGCACTAGCACTAGCTTCAAGGTTCCATGCA---AACGGTGT 2152
Qy 621 SerLeuThrLeuGlyLeuProSerGluAsnLeuAlaMetProValSerGlnGlnAsn 640
Db 2153 TCTCTCACACTTGGACTTCCCAATGAGAACCTTCTATGTCAGGAACCTCAACACAT 2212
Qy 641 ---TyrLeuSerAsnAsn-----LeuGlySerArgSerGluMetGlySerHisTyrAsn 657
Db 2213 GGTTCTCTCACAGATATACACCTGGGAGGAGACTTGAATGGACAAATGGNAAT 2272
Qy 658 ArgMet-----GlyTyrGluAsn 663
Db 2273 GAGTTTGTACTATCAACACTGCTCCACCTTCTTCTCACTCAGGAACCACTTATGAGAGC 2332
Qy 664 IleAspPheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
Db 2333 ATTGATATTCAAAACAGAAAGAGGTTGTTGCTCATCAGTTATTACAGATTTTCGTG 2389

RESULT 2
US-10-425-114-13353
; Sequence 13353, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13353
; LENGTH: 1555
; TYPE: DNA
; FEATURE:
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: JC-QMST02400042B10_F1J
US-10-425-114-13353

Alignment Scores:
Pred. No.: 6,33e-107 Length: 1555
Score: 1174.50 Matches: 271
Percent Similarity: 66.06% Conservative: 58
Best Local Similarity: 54.42% Mismatches: 90
Query Match: 32.96% Indels: 79
DB: 18 Gaps: 20

US-10-624-201A-2 (1-688) x US-10-425-114-13353 (1-1555)
Qy 240 AsnLysGluSer-----MetProLeuAlaSerAspValAsnThrAsnSerSerGlyGly 257
Db 5 AATAGGAATCAACTAATCTGGTGGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 64
Qy 258 GlyCysSerSerArgGlnLysAsnGluValAlaValGluLeuThrAlaGlnArg 277
Db 65 GGAGAAACACAGTCAGGGAACAA-----GTGGTTCAACTCAGCACCGCTCAGAG 115
Qy 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluValGluGlnArg 297
Db 116 CAAGAGCTTCAGATGAAGAAGTCCAGCTTGTGACCATGCTCGATGAGGTAGAACAAAG 175
```

```

Qy 298 TyrArgGlnTyrHisGlnMetGlnIleValLeuSerPheGluGlnValAlaGly 317
Db 176 TACAGACAATATCACCACCAATGCAAAATTGGTATCGTCAATTGAGCAGCAGCAGGT 235
Qy 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db 236 TATGGGGCGGCAAAATCTTACACTGCCCTTAAAGACTATCTCAAAAGCAATTTCAGG 295
Qy 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 296 TGCCTTAAAGATGCATCTCTGCACAAATCAAAAGCCACGACGACATTTGGTGAAGAT 355
Qy 358 GluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisLeuArg 377
Db 356 GATTGTTTAGGAGTTAAAGTAGAAGGTCGAGGCTTAGGTTTGTTCACCACTCTCAGG 415
Qy 378 GlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTyrArgProGlnArg 397
Db 416 CAACAAGGGCACTTCAGCAGCTAGGAATGATCAACCCCAATCATGGAGGCCCAAGA 475
Qy 398 GlyLeuProGluArgAlaValSerValLeuArgAlaTyrLeuPheGluHisPheLeuHis 417
Db 476 GGTTCGCTGACGAGCTGTTCCATCTTCGGGCTTGGCTTTTGGAGCATTTCTTCAC 535
Qy 418 ProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSer 437
Db 536 CCTTATCCAAAGGACTCCGATAAGTTATGCTTGTCTAAACAAACTGGACTTGTAGGAGC 595
Qy 438 GlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGlu 457
Db 596 CAGGTGTCAAACTGGTTTATCAATCCCGAGTTTCGGCTTGGAGGCCAATGGTTGAAGNA 655
Qy 458 MetTyrLeuGluGluValLysAsnGlnGluGln-----AsnSerThrAsnThrSerGly 475
Db 656 ATGTACTTGGAGAAATCAAGGAGCATGAACAGGTAATGGCTCAGAAAACACCAATCA 715
Qy 476 AspAsnLysAsnLysGlu-----ThrAsnIleSerAlaProAsn-----GluGlu 490
Db 716 AAAGAATCGAGCAAGGAGTTGGCTTCAACAGACAAATGTTGCACATCATCTCCAATCA 775
Qy 491 LysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGlu 510
Db 776 AAACATGAAAGCTTCAACAAACCA-----AATACTTCCCCAACCGAG 817
Qy 511 IleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHisAlaHis 530
Db 818 ATCTCAACAGCTCT---TCATGTCTCCAATGGGAGGTCCTTCAATCCCACTCT--- 871
Qy 531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle 550
Db 872 GGTTTCCATCTTGCAGGATCATCTGACATG-----CAAAT 907
Qy 551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
Db 908 AGGAGTCCAAACAAACCAAGAGTTCTGAAATGCAAAAC---TCCCAAGTAGTATCTCT 964
Qy 571 SerSerValAspMetGluAlaLys-----AlaArgGluSerSer 583
Db 965 ---TCAGTGGACATGGAATGAAGCATAGTGGTGACCATGGAATATAGAGATGCCAAC 1021
Qy 584 AsnLys-----GlyPheThrAsnProLeuMet 592
Db 1022 ACAAGTTTGGTATTGAAAGGCATCACCACAAAGATGGGTAC-----CCTTTGATG 1075
Qy 593 -----AlaAlaTyrAlaMetGlyAspPheGly--- 601
Db 1076 ACTAGCAATCCAAACCATGGTGGGGGTTGGAGCATTCACCATGGAAGCATTTGGAAGC 1135
Qy 602 ArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHisGlyAsnAenglyValSer 621
Db 1136 AGATTAAATGCTACTACTAGGCAACTAGCTTCAAGGTTCCATGCA---AACGGTGTTCCT 1192
```

```
QY 622 LeuThrLeuGlyLeuProSerGluAsnLeuAlaMetProValSerGlnGlnAsn--- 640
Db 1193 CTCACATTTGGACTTCCCACTGAGAACCTTTCTATGTCAGGAACCTCAAACAATGGG 1252
QY 641 TyrLeuSerAsnAsp-----LeuGlySerArgSerGluMetGlySerHisTyrAsnArg 658
Db 1253 TTCTCTCACAGAAATATACACCTGGGAAGGAGACTTGAATATGGAAACAAATGGAATGAG 1312
QY 659 Met-----GlyTyrGluAsnIle 664
Db 1313 TTTTGTACTATCAACACTGCTCACCTTTCTCTACTCAGGAACCACTTATGAGACATT 1372
QY 665 AspPheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
Db 1373 GATATTCAAAACAGAAGAGGTTTGTGCTCATCAGTTATTACCAGATTTCGTG 1426

RESULT 3
US-09-938-842A-1337
; Sequence 1337, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1337
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1337

Alignment Scores:
Pred. No.: 2 4e-104 Length: 2043
Score: 1150.50 Matches: 306
Percent Similarity: 51.81% Conservative: 81
Best Local Similarity: 40.96% Mismatches: 201
Query Match: 32.29% Indels: 159
DB: 9 Gaps: 28

US-10-624-201A-2 (1-688) x US-09-938-842A-1337 (1-2043)
QY 22 HisGlyAsn-----SerAsnAsnAsnIleGlnThrLeuTyrLeu 35
Db 16 CACGGAACCCACCGGAGACTCTGCGCGATCGACGGTGTCTTCAACGTTGATCCTC 75
QY 36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGln 55
Db 76 ATGAATCCAACTACTTACGTTTCAG---TACACCCAAACAAGACAACGACTCGAACACAAC 132
QY 56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCyshHisAlaAsnIle--- 74
Db 133 -----AACAAAGCAACATAGCAACACAAACAACAACAACAACAACAAC 180
QY 75 -----GlnHisAlaProLeu-----GlnGlnGln 82
Db 181 AACAAACAACAGTAGTTTCGTTTTCCTCGATTCCACGCGCGAGCAAAACGCGAGCCAG 240
QY 83 HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97
Db 241 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCATATACAGCCCGCGAACAAC 300
```

```
QY 98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
Db 301 ATCTCGGTACTTTCACGGTTATCTCTCGCGCGTGCAGTACAGTCTCTACGGTAGCCACAA 360
QY 117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
Db 361 GTG----- 363
QY 137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154
Db 364 -----GATCCCACTCACCAAGCCGCG 387
QY 155 -----GlnGlnGlnGlnGlyLeuSerLeuSerProGlnLeuGlnGln 172
Db 388 TGTGAGAGCCACGCGCGCAGCAAGCCCTCTCTTTAACCTCTCTGCTCTCAACAGAGCAG 447
QY 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerSerProArgThr 187
Db 448 CAACAGCAACATCATCAACCAACACAGCCCTATTACGTCGGATTCCGGTCCGAGCATGGA 507
QY 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202
Db 508 GAAGATATCCGGTTCGGTCTTACAGGATCGGGGTAAACAACGGTATAGTAAT 567
QY 203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
Db 568 CTTGTT---AGCTCCAAGTACTTTGAAGCAGACACAAGAGCTTCTTGACGAAGTAGTCAAC 624
QY 223 IleValGlyLysSerIleLysGlyAspGln-----LysLysAspAsnSer 238
Db 625 GCTGATTCCGATGATCAATCAACGCTAAATCCCACTATTCTCATCGAAAAAGGTAGTTGC 684
QY 239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
Db 685 GGAAATGATAAACCTGTCGAGAAATCATCGCGCGCGCTGGAGGAGAGGTTCGCGTGC 744
QY 258 GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db 745 GGAGCAGAAGCAGCGCGGNAACGT-----CCGTTGGAGCTAGGCACGCGCAGAGAGA 795
QY 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGlnArg 297
Db 796 CAAGAAATACAGATGAAGAAAGCAAACTTAGTAACATGCTTTCATCAGGTGAGCAGAGA 855
QY 298 TyrArgGlnTyrHisGlnMetGlnIleValLeuSerPheGlnValAlaGly 317
Db 856 TATAGACAGTACCACAGCAGATGCGAGTGGTCTCTTCGTTGAGCAACGCGCAGG 915
QY 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db 916 ATAGGATCAGCGAAGTCTACACGTCGTCAGTATTGAAAAACCATATCAAGACAGTTCGCT 975
QY 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGluGlu 357
Db 976 TGTCTTAAAGAGCGCATCGTGTGAGTAAAGCGGCCAACAAAGAGTCTTGGGAGGAA 1035
QY 358 GlnGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
Db 1036 GATTCAAGTCTCGGTGTGGGAGGTTTGAGGGGTGAGGTCGAGTTCGTTGAGCCAC 1095
QY 376 LeuArgGlnGlnArgAlaLeuGlnIleGlyMetMetGln-----ProAsnAla 392
Db 1096 TTGAGACAGCAAGAGCTCTTCAACACTGGGAATGATTCAACATCCTTCCCAATATGCT 1155
QY 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db 1156 TGGAGACCTCAACGCGGTCTCCAGAACGAGCCGCTCAGTTCCTCGTCTTGGCTTTC 1215
QY 413 GlnHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 1216 GAACACTTCTTCATCCATACCTAAGGATTCGGAACAGCAGCATGCTAGTAGCAACA 1275
QY 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
```


Db 625 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGCTAGTTGC 684
Qy 239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
Db 685 GGAAATGATAAACCCTGTCGAGAAATCATCGCGCGCTGGAGGAGAAAGTTCCCGTGGC 744
Qy 258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db 745 GGAGCAGAAAGCAGCCGGAAACGT-----CCGGTGGAGCTAGGCACGCGAGAGAGA 795
Qy 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluValGluGlnArg 297
Db 796 CAAGAATAACAGATGAAGAAGCAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 855
Qy 298 TyrArgGlnThrHisGlnMetGlnIleLeValLeuSerPheGluGlnValAlaGly 317
Db 856 TATAGACAGTACCACGACGATGACAGATGGTGTCTCTTCGTTCGAGCAAGCGGCGAGG 915
Qy 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db 916 ATAGGATCAGCGAAAGTCATACAGCTCGCTAGCATTTGAAACCATATCAAGACAGATTCCT 975
Qy 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 976 TGCCTTGAAGAGCGGATCGCTGTCAGATAAAGCGCCCAACAAGAGAGTCTTGGGAGGAA 1035
Qy 358 GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHis 375
Db 1036 GATTCAGTGTCTGGTGTGGAGGTTTGAGGGTTCGAGGCTCAAGTTCGTGGACCAACCAC 1095
Qy 376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla 392
Db 1096 TTGAGACAGCAAGAGAGCTCTCAACAACATGGGAATGATTCACATCCTCCCAATATGCT 1155
Qy 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db 1156 TGGAGACCTCAACGTGGTCTCCAGAACGAGCGCTCTCAGTTCTCGTCTGGCTCTTC 1215
Qy 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 1216 GAACATTTCTTCATCCATCCCTAAGGATTCGACACACATGCTAGCTAAGCAACA 1275
Qy 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db 1276 GGACTCACTCGTAGCCAGGTGTCGAAGTGTATTAACGCGAGAGTTCGGTTATGAAA 1335
Qy 453 ProMetValGluGluMetTyrLeuGluValLysAsnGlnGluGlnAsnSerThrAsn 472
Db 1336 CCAATGGTGGAGAGATGTACATGGAGGAATGAAGGAGCAGGCAAGAACAATG----- 1389
Qy 473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
Db 1390 -----GGATCCATGGAAAGAGATCCTTTGGATCAAAAGCAACGAAGAT----- 1431
Qy 493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSer 512
Db 1432 -----TCTGCTTCAAAGTCA 1446
Qy 513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis 530
Db 1447 ACAAGTAAACAAGAAAGAGCCCAATGGGGGACACTAATATACCATATGAATCCCAATCAC 1506
Qy 531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle 550
Db 1507 AACGGTGACTAGAACGGCTCACTGGAATGCAAGGAAGCCCCMAGAGACTAAGAACACAC 1566
Qy 551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
Db 1567 GACGACACAATGATGCAGCAATAAAT-----GCGGATTCAGCTCCCAACGAGAGACTC 1620
Qy 571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db 1621 ACGATGAAGATTCTAGAAAGACGGCAAGGGATAAGATCAGATGGTGGCTAC-----CCT 1674

Qy 591 LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db 1675 TTCATGGGTAATTTCCGGCAATCAAAATGGATGAGATGTCGAAGATTTGATGTAGTCTCA 1734
Qy 607 AspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLeuThrLeu 624
Db 1735 GACCAGGAGCTCATGGCGAAAGGTACTCAGAGAAACAACAATGGCGTGTCCCTCAGTTA 1794
Qy 625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln-----GlnAsnTyr 641
Db 1795 GGTTCACCTCATTTGATAGCTTGTCCACGACCATCAGGGTTTCATGCAGACCCAC 1854
Qy 642 LeuSerAsnAspLeuGlySerArgSerGluMetGly----- 653
Db 1855 CATGGGATTCCTATAGGAGAGAGTGAATAAGAGAAACAGAGAAATATGGACCCGCC 1914
Qy 654 -----SerHis 655
Db 1915 ACCATCAATGGTGTAGTTCGACCAACCGCACATTCATCAGCGCAGCTGCCGCGCT 1974
Qy 656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675
Db 1975 TACAATGGGATGAACATACAGAAC-----CAGAAGAGATATGTGGCT 2016
Qy 676 GlnLeuLeuProAspPheVal 682
Db 2017 CAGTTATTGCCCGCACTTCGTT 2037

RESULT 5

US-10-495-918-153
; Sequence 153, Application US/10495918
; Publication No. US20050009187A1
; GENERAL INFORMATION:
; APPLICANT: RIKEN
; TITLE OF INVENTION: A STRESS-RESPONSIVE PROMOTER AND A GENE CODING FOR A STRESS-RESP
; FILE REFERENCE: RFH14-024T
; CURRENT APPLICATION NUMBER: US/10/495,918
; PRIOR FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: JP 2001-353038
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: JP 2002-20329
; PRIOR FILING DATE: 2002-01-29
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 153
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2040)
US-10-495-918-153

Alignment Scores:
Pred. No.: 2,4e-104 Length: 2043
Score: 1150.50 Matches: 306
Percent Similarity: 51.81% Conservative: 81
Best Local Similarity: 40.96% Mismatches: 201
Query Match: 32.23% Indels: 159
DB: 21 Gaps: 28

US-10-624-201A-2 (1-688) x US-10-495-918-153 (1-2043)

Qy 22 HisGlyAsn-----SerAsnAsnAsnIleGlnThrLeuTyrLeu 35
Db 16 CACGGAACCCACCGAGATCTCTCGCGATCCGACGGTGGTCTTCAAAACGTTGATCCTC 75
Qy 36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln 55
Db 76 ATGAATCCAACCTACTTACGTTTCAG---TACCCCAACAGACAGACCTCGAACAACAAC 132

[illegible]

Db	1096	TTGAGACAGCAAGAGCTCTTCAACCACTCGGGAATGATTCAACATCCTTCCAATAATGCT	1155
Qy	393	TriArgProGlnArgGlyLeuProGluuAqAlaValSerValLeuArgAlaTprLeuPhe	412
Db	1156	TGGAGACCTCAACGTGTCTCCACAGACGCGTCTCAGTTCTCCGTGTGGCTTCTTC	1215
Qy	413	GluHisPheLeuHisProTyrProLysAspSerAspLysileMetLeuAlaLysGlnThr	432
Db	1216	GAACACATTTCTTCATCCATACCCCTAAGATTTCGACAAACACATGCTAGCTAAGCAACA	1275
Qy	433	GlyLeuThrArgSerGlnValSerAsnTprPheIleAsnAlaArgValArgLeuTprLys	452
Db	1276	GGACTCACTCGTAGCCAGGTGTGCAACTGTGTATATAACCGCAGAGTTCGGTTATGGAAA	1335
Qy	453	ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn	472
Db	1336	CCAAATGTGAGGAGATGTACATGGAGGAATGAAGGACGCGCAAGAACAATG	1389
Qy	473	ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis	492
Db	1390	-----GGATCCATGGAAAAGACTCTTTGGATCAAAACCAACGAAGAT-----	1431
Qy	493	ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer	512
Db	1432	-----TCTGCTTCAAAAGTCA	1446
Qy	513	ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis	530
Db	1447	ACAAGTAACCAAGAAAGACCCCAATGGCGGACACTAATTAACCATATGAATCCCAATCAC	1506
Qy	531	AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle	550
Db	1507	AACGGTGACCTAGNAGGCGTCACTGGGAATCAAGGAAGCCCAAGAGACTAAGAACCCAGC	1566
Qy	551	GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu	570
Db	1567	GACGAGACAATGATGACGCCCAATAAAT-----CGCGATTTCAGTCTCAACGAGAAGCTC	1620
Qy	571	SerSerValAspMetGluAlaLysAlaArgLysSerSerAsnLysGlyPheThrAsnPro	590
Db	1621	ACGATGAAGATTCTTAGAAGAACGCGCAAGGATGAAGATCAGATGGTGGCTAC-----	1674
Qy	591	LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis	606
Db	1675	TTCATGGGTAATTTCGGGCAATACCAATGGATGAGATGTCAAGATTGTGATAGTCTCA	1734
Qy	607	AspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLeuThrLeu	624
Db	1735	GACCAGAGCTCATGGCGCAAGGTACTCAGGAAACAACAATGCGGTGTCCCTCAGCTTA	1794
Qy	625	GlyLeuProSerGluAsnLeuAlaMetProValSerGln-----GlnAsnTyr	641
Db	1795	GGTTTACCTCATGTGTAGTCTGTGTCACGCAACCATCAGGGTTTCATGCAGACCCAC	1854
Qy	642	LeuSerAsnAspLeuGlySerArgSerGluMetGly-----	653
Db	1855	CATGGGATTCTATAGGAGAAAGTAGTCAAAATAGGAGAAACACAGGNAATATGACCGGCC	1914
Qy	654	-----SerHis	655
Db	1915	ACCATCAATGGTGTAGTCTGCACCAACCGCACATTCATCAGCGGCAGCTGCGCGGCT	1974
Qy	656	TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr	675
Db	1975	TACAATCGGATGAACATACAGAAC-----CAGAAGAGATATGTGGCT	2016
Qy	676	GlnLeuLeuProAspPheVal	682
Db	2017	CAGTTATGCCGACATTCGTT	2037

RESULT 6

US-10-225-066A-335

```
; Sequence 335, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBIO036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; PRIOR FILING DATE: 2002-08-09
; PRIOR FILING DATE: 2001-08-09
; PRIOR FILING DATE: 2001-04-18
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 335
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-225-066A-335

Alignment Scores:
Pred. No.: 4,81e-104 Length: 2385
Score: 1148.50 Matches: 306
Percent Similarity: 51.67% Conservative: 80
Best Local Similarity: 40.96% Mismatches: 202
Query Match: 32.23% Indels: 159
DB: 17 Gaps: 28

US-10-624-201a-2 (1-688) x US-10-225-066A-335 (1-2385)
QY 22 HisGlyAsn-----SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
Db 194 CACGGAACCCACCGGAGATCTCGCGGATCCGACGGTGTCTTCAAACGTTGATCCTC 253
QY 36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGln 55
Db 254 ATGAATCCCAACTACTTACGTTTCAG-----TACACCCCAACAAGACACGACTCGAACACAC 310
QY 56 LeuLeuPheLeuAsnSerProAlaAlaSerAsnAlaLeuCyehHisAlaAsnIle--- 74
Db 311 -----AACACACGACCAATACCAACACACACACACACACACACACACACAC 358
QY 75 -----GlnHisAlaProLeu-----GlnGlnGln 82
Db 359 AACACACACAGTAGTTTCGTTTTCCTCGATTCCACGCGCGCAGCAACACGAGCCAG 418
QY 83 HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97
Db 419 CAGTTCGTCGGAATACCACTCTCAGGTCACGAGCTGCTTCCATTACAGCCCGCGACAC 478
QY 98 IleAsn---HisHisGlyLeuLeuGlnArgMetTyrAsnAsnGlnAspGlnSerGln 116
Db 479 ATCTCGGTACTTCACGGTTATCTCTCCGCGGTGTCAGTACAGTCTCTACGTTAGCCACAA 538
QY 117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
; 539 GTG----- 541
; 137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154
; 542 -----GATCCCACTCACCAAGCCGCG 565
; 155 -----GlnGlnGlnGlnGlnGlyLeuSerLeuSerProGlnLeuGlnGln 172
; 566 TGTGAGACCGCAGCGCGCAGCAAGCCCTCTTTAAACCTCTCTCTCAACAGCAGCAG 625
; 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerSerProArgThr 187
; 626 CAACAGCAACATCATCAACACACAGCCTATTACGTCGGATTCCGGTCCGACATCGA 685
; 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202
; 686 GAAGATATCCCGGTCGGGTCTGCTCTACAGATCGGGGTAAACAAACGGTATAGCTAAT 745
; 203 MetValLeuGlySerTyrLeuLysAlaAlaGlnGlnLeuLeuAspGluValValAsn 222
; 746 CTTGTT---AGCTCCAAGTACTTGAAGCAGCAGCAGAGCTTCTTGACGAAGTAGTCAAC 802
; 223 IleValGlyLysSerIleLysGlyAspAspGln-----LysLysAspAsnSer 238
; 803 GCTGATTCCGATGATGAACGCTAAATCCCACTATTCTCATCGAAAAAGGTAGTTGC 862
; 239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
; 863 GGAATATGATAAACCTGTCGAGAAATCATCGGCGCGCTGGAGAGAAAGTTCCCGTGC 922
; 258 GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
; 923 GGAGCAGAGCAGCGCGGAAACGT-----CCGTTGGAGCTAGGCACGCGCAGAGAGA 973
; 278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
; 974 CAAGAAATACAGATGAAGAAAGCAAACTTAGTAACATGCTTTCATGAGGTGGCAGAGA 1033
; 298 TyrArgGlnTyrHisGlnMetGlnIleValLeuSerPheGlnGlnValAlaGly 317
; 1034 TATAGACAGTACCACGACGATGAGATGGTGATCTCTTCGTTGAGCAACGCGGAGGG 1093
; 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
; 1094 ATAGGATCAGCGAAGTCATACAGTCGCTAGCATTTGAAACCCATATCAAGACAGTCCGT 1153
; 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGluGlu 357
; 1154 TGCTTGAAGAGCGATCGCTGTCAGATAAAAGCGGCCCAACAGAGTCTTGGGAGGAA 1213
; 358 GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHis 375
; 1214 GATTCAAGTCTCGGTGTTGGGAGGTTTGAAGGTCGAGGCTCAAGTTCGTGACCAACC 1273
; 376 LeuArgGlnGlnArgAlaLeuGlnIleGlyMetMetGln-----ProAsnAla 392
; 1274 TTGAGACAGCAAGAGCTCTTCAACACTGGGAATGATTCAACATCCTTCCAATATGCT 1333
; 393 TyrArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTyrLeuPhe 412
; 1334 TGGAGACCTCAACGTTGTCCTCCAGAACGAGCGCTCTCAGTTCTCGTCTTGGCTCTTC 1393
; 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
; 1394 GAACACTTCTTTCATCATACCTCAAGGATTCGGAACAGCATGCTAGCTAAGCAACA 1453
; 433 GlyLeuThrArgSerGlnValSerAsnTyrPheIleAsnAlaArgValArgLeuTyrLys 452
; 1454 GGACTCACTCGTAGCCAGGTGTCGAACCTGGTTTATAAACCGCAGAGATTCGGTTATGAAA 1513
; 453 ProMetValGluGluMetTyrLeuGluValLysAsnGlnGluGlnAsnSerThrAsn 472
```


Db 542 -----GATCCCACTCACCAGCAAGCGCG 565
Qy 155 -----GlnGlnGlnGlnGlyLeuSerLeuSerLeuSerProGlnLeuGln 172
Db 566 TGTGAGACGCCAGCGCGCAGCAAGCCTCTCTTTAAACCTCTCGTCTCAACAGCAGCAG 625
Qy 173 GlnLeuSerPheAsnAsn-----AsnLeuSerSerSerProArgThr 187
Db 626 CAACACCAATCATCAACACACACAGCCTATTCAGTCGGATTCGGGTCCGGACATGGA 685
Qy 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202
Db 686 GAAGATATCCGGTCCGGTCTGCTCTACAGGATCGGGGTAAACAAACGCTATAGCTAAT 745
Qy 203 MetValLeuGlySerIleValLeuAlaGlnGlnLeuLeuAspGluValValAsn 222
Db 746 CTTGTT---AGTCCCAAGTACTTGAAGGAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 802
Qy 223 IleValGlyLeuSerIleValLeuGlyAspGln-----LysLeuAspAsnSer 238
Db 803 GCTGATTCGATGCATGACGCTAAATCCCACTATTCTCATCGAAAGGGTAGTTGC 862
Qy 239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
Db 863 GGAATATGATAAACCTGTGCGAGAATCATCGCGCGCGCTGGAGGAAAGGTTCCGGTGGC 922
Qy 258 GlyLeuSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db 923 GGAGCAGAGCAGCGCGGAAACGT-----CCGTGGAGCTAGGCACGCGCAGAGAGA 973
Qy 278 GlnGluLeuGlnMetLysAlaLysLeuLeuAlaMetLeuGluValGluGlnArg 297
Db 974 CAAGAATACAGATGAAGAAGCAAACTTAGTAAATGCTTATGAGGTGGAGCAGAGA 1033
Qy 298 TyrArgGlnThrHisGlnMetGlnIleValLeuSerPheGluGlnValAlaGly 317
Db 1034 TATAGACAGTACCACCAGCAGATGCAGATGCTCTTCGTTCCAGCAAGCGCGCAGG 1093
Qy 318 IleGlySerAlaLysSerThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db 1094 ATAGGATCAGCGAAGTCTATACACGTCGCTAGCATTTGAAACCATATCAAGACAGTTCCT 1153
Qy 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 1154 TGTCTGAAGAGCGATCGCTGTCAGATAAAGCGCGCCCAACAGAGTCTTGGGAGGAA 1213
Qy 358 GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHis 375
Db 1214 GATTCAGTGTCTGCTGTGGAGGTTTGAAGGCTCAGGCTCAAGTTCGTGGACCAACCAC 1273
Qy 376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln-----ProAsnAla 392
Db 1274 TTGAGACAGCAAGAGCTCTTCAACAACTGGGAATGATTCACATCCTTCCCAATAATGCT 1333
Qy 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db 1334 TGGAGACCTCAAGTGTCTCCAGAACGAGCGCTCTCAGTCTCGTCTGCTTGGCTCTTC 1393
Qy 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 1394 GAACACTTTCTTCATCCATACCTCAAGGATTCGCAAGCAGACATGCTAGCTAAGCAACA 1453
Qy 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db 1454 GGACTCAGTCGTAGCCAGGCTCGAACTGTTTATAACCGAGAGTTCGGTTATGGAA 1513
Qy 453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGlnAsnSerThrAsn 472
Db 1514 CCAATGTTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGCAAGAACATG----- 1567
Qy 473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
Db 1568 -----GGATCCATGGAAAGACTCTTTGGATCAAAAGCAACGAAGAT----- 1609

Qy 493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSer 512
Db 1610 -----TCTGCTTCAAAGTCA 1624
Qy 513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis 530
Db 1625 ACAAGTAACCAAGAAAAGCCCAATGCGGACACTAATTACCATATGATCCCATCAC 1684
Qy 531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle 550
Db 1685 AACCGTGACCTAGAAGCGCTCACTGGAATGCAAGGATGCCCAAGAGACTAAGAACCAGC 1744
Qy 551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerIleLeu 570
Db 1745 GACGACAGCAATGATGAGCCCAATAAAT-----GCGGATTTCACTCCACAGAAAGCTC 1798
Qy 571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db 1799 ACGATGAAGATTCTAGAGAACCGCAAGGATAAGATCAGATCGTGGCTAC-----CCT 1852
Qy 591 LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db 1853 TTCAATGGGTAAATTTCCGGCAATACCAATGAGATGATGCAAGATTGATGTAGTCTCA 1912
Qy 607 AspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLeuThrLeu 624
Db 1913 GACGAGGACTCATGGCGCAAGGTACTCAGGAACAACAATGGCGTGCTCCCTCAGCTTA 1972
Qy 625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln-----GlnAsnTyr 641
Db 1973 GGTTTACCTCATTTGATAGTGTTCGTCACCGACCATCAGGGTTTCATGCAGACCCAC 2032
Qy 642 LeuSerAsnAspLeuGlySerArgSerGluMetGly-----SerHis 655
Db 2033 CATGGGATTCCTATAGGAGAGAGTGAATAATAGGAGAAACAGAGGAATATGGACCCGCC 2092
Qy 654 -----SerHis 655
Db 2093 ACATCAATGGTGGTAGTCTGCACCAACCGCACATTCATCAGCGCAGCTGCCCGGCT 2152
Qy 656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675
Db 2153 TACAATGGGATGAACATACAGAAC-----CAGAAAGAGATATGTGGCT 2194
Qy 676 GlnLeuLeuProAspPheVal 682
Db 2195 CAGTTATTGCCCGACTTCGTT 2215

RESULT 8

US-10-225-066A-335
; Sequence 335, Application US/10225066A
; Publication No. US20050160493A9
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omalra
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225, 066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18

;; PRIOR APPLICATION NUMBER: 60/310,847
;; PRIOR FILING DATE: 2001-08-09
;; PRIOR APPLICATION NUMBER: 60/336,049
;; PRIOR FILING DATE: 2001-12-05
;; PRIOR APPLICATION NUMBER: 60/338,692
;; PRIOR FILING DATE: 2001-12-11
;; PRIOR APPLICATION NUMBER: 10/171,468
;; PRIOR FILING DATE: 2002-06-14
;; NUMBER OF SEQ ID NOS: 1122
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 335
;; LENGTH: 2385
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-10-225-066A-335

Alignment Scores:

Pred. No.: 4,81e-104 Length: 2385
Score: 1148.50 Matches: 306
Percent Similarity: 51.67% Conservative: 80
Best Local Similarity: 40.96% Mismatches: 202
Query Match: 32.23% Indels: 159
DB: 22 Gaps: 28

US-10-624-201A-2 (1-688) x US-10-225-066A-335 (1-2385)

Qy 22 HisGlyAsn-----SerAsnAsnAsnAsnGlnThrLeuTyrLeu 35
Db 194 CACGAAACCCACCGGAGACTCTGCGCGATCCGACGGTGGTCTTCAACGTTGATCCTC 253
Qy 36 MetAsnProAsnAenTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGln 55
Db 254 ATGAATCCAACTACTTACGTTTCAG---TACACCAACACAGACACGACTCGAACACAC 310
Qy 56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
Db 311 -----AACACACGACAAATAGCAACACACACACACACACACACACAC 358
Qy 75 -----GlnHisAlaProLeu-----GlnGlnGln 82
Db 359 AACACACACAGTAGTTTCGTTTCTCGATTCACACGCGCGGACCAACGCGGACCG 418
Qy 83 HisPheValGlyValProLeu-----ProAlaValSerLeuHisAspGln 97
Db 419 CAGTTCGTCGGAATACCACTCTCAGGTTCACGAAGACTCTTCCATTACAGCGCGGAC 478
Qy 98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGln 116
Db 479 ATCTCCGTACTTCACGGTTATCTCTCCGCGGTGCAGTACAGTCTCTACGGTAGCCCAA 538
Qy 117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
Db 539 GTG----- 541
Qy 137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln----- 154
Db 542 -----GATCCCACTCACACGACCAACGCGCG 565
Qy 155 -----GlnGlnGlnGlnGlnGlyLeuSerLeuSerLeuSerProGlnLeuGln 172
Db 566 TGTGAGACGCCGCGCGGACCAAGGCTCTCTTTAACCTCTCGTCTCAACGACGACG 625
Qy 173 GlnIleSerPheAsnAsn-----AsnIleSerSerSerProArgThr 187
Db 626 CAACAGCAACATCATCAACACACACGCTTATTCACGTCGGATTCGGGTCGGACATGGA 685
Qy 188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer-----SerAsn 202
Db 686 GAAGATATCGGGTGGGCTCTGGCTCTACAGGATCGGGGTAAACAAACGCTATAGCTAAT 745
Qy 203 MetValLeuGlySerIleTyrLeuLeysAlaAlaGlnIleLeuLeuAspGluValValAsn 222
Db 746 CTTGTT---AGCTCCAAGTACTTGAAGGCGACACAGAGCTTCTTGACGAAGTAGTCAAC 802

Qy 223 IleValGlyLeysSerIleLeysGlyAspAspGln-----LysLysAspAsnSer 238
Db 803 GCTGATTCGATGACATGAACGCTAAATCCCACTATTCTCATGAAAGGGTAGTTCG 862
Qy 239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
Db 863 GGAATGATAAACCTGTCGGAGAATCATCGCGCGCGCTCGAGGAGAAGGTTCCGCTGGC 922
Qy 258 GlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrAlaGlnArg 277
Db 923 GGAGCAGACGACCGCGGAACGCT-----CCGGTGGAGCTAGGACGCGCAGAGAGA 973
Qy 278 GlnGluLeuGlnMetLysLysAlaLysLeuAlaMetLeuGluGluValGluGlnArg 297
Db 974 CAAGAAATACAGATGAAGAAAGCAAACTTAGTAAACATGCTTCATGAGGTGGAGCAGAGA 1033
Qy 298 TyrArgGlnTyrHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
Db 1034 TATAGACAGTACCACCGACGAGATGCGATGGTGTATCTTCTGTTCCGACGAAGCGCAGGG 1093
Qy 318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db 1094 ATAGGATCAGCGAAGTCATACGCTCGCTAGCATTTGAAACCATATCAACACAGTTCCTG 1153
Qy 338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db 1154 TGCTTGAAGAGCGGATCGCTGCTGCTAGATAAAGCGGCCAACAGAGTCTTGGGGAGAA 1213
Qy 358 GluGlyLeuGly-----GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
Db 1214 GATTAGTGTCTGGTGTGGAGGTTTGGAGGCTCGAGGCTCAAGTTCTGCGGACCCAC 1273
Qy 376 LeuArgGlnGlnArgAlaLeuGlnIleGlyMetMetGln-----ProAsnAla 392
Db 1274 TTGAGACAGCAAGAGCTCTTCAACACCTGGGAATGATTCAACATCTCTCAATAATGCT 1333
Qy 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db 1334 TGGAGACCTCAACGTGCTCCCGACCAACGAGCGCTCTCAGTCTCTCGTCTGCTCTTC 1393
Qy 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 1394 GAACACTTTCTTCATCATACCTCAAGGATTCGACAAAGCACATGCTAGCTAAGCAACA 1453
Qy 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db 1454 GGACTCAGTCTGAGCGGCTCGAACTGGTTTATAACCGGAGAGTTCGGTTATGGAA 1513
Qy 453 ProMetValGluGluMetTyrLeuGluValLysAsnGlnGluGlnAsnSerThrAsn 472
Db 1514 CCAATGTTGGAGGAGATGTACATGGAGGAAATCAGGAGCAGCAAGAACATG----- 1567
Qy 473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
Db 1568 -----GGATCCATCGAAAGACTCTTTGGATCAAAAGCAACGAAGAT----- 1609
Qy 493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSer 512
Db 1610 -----TCTGCTTCAAAAGTCA 1624
Qy 513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis-----HisAlaHis 530
Db 1625 ACAAGTAACCAAGAAAGACCCCAATGCGGACACTAATTTACCATATGAATCCCAATC 1684
Qy 531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle 550
Db 1685 AACGGTGCCTTAGAGCGCTCACTGGAATGCAAGGATGCCCCAAGAGACTAAGAACCCAGC 1744
Qy 551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
Db 1745 GACGACGAATGATGACGACCAATAAAT-----GCGGATTTAGCTTCAACGAGAGCTC 1798

```
QY 571 SerSerValaspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db 1799 ACATGAAGATTCAGAGAACCGCAGGAGTAAGATCAGATGGTGGCTAC-----CCT 1852
QY 591 LeuMet-----AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db 1853 TTCATGGGTAAATTCGGGCAATACCAAAATGGATGAGATGTCGAAGATTTGATGTAGTCTCA 1912
QY 607 AspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSerLeuThrLeu 624
Db 1913 GACCAGGAGCTCATGGCGCAAGAGTACTCAGGAACAACAATGGCGTGTCCCTCACGTTA 1972
QY 625 GlyLeuProSerGluAsnLeuAlaMetProValSerGln-----GlnAsnTyr 641
Db 1973 GGTTHACCTCATGTGATAGCTGTGCTCCACGACCATCAGGGTTTCATGCAGACCCAC 2032
QY 642 LeuSerAsnAspLeuGlySerArgSerGluMetGly-----SerHis 655
Db 2033 CATGGGATTCCTATAGCGAGAGAGTGAAATAGGAGAAACAGAGGAATATGGACCCGCC 2092
QY 654 -----SerHis 655
Db 2093 ACCATCAATGGTGTAGCTCGACCAACAACCGCACATTCATCAGCGGCGAGCTGCCGGCT 2152
QY 656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675
Db 2153 TACAATGGGATGAACATACAGAAC-----CAGAGAGATATGTGGCT 2194
QY 676 GlnLeuLeuProAspPheVal 682
Db 2195 CAGTTATGCCCGACTTCGTT 2215

RESULT 9
US-10-424-599-133558
; Sequence 133558, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 133558
; LENGTH: 1822
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91608C.1
US-10-424-599-133558

Alignment Scores:
Pred. No.: 2 58e-97 Length: 1822
Score: 1079.50 Matches: 248
Percent Similarity: 64.19% Conservative: 55
Best Local Similarity: 52.54% Mismatches: 112
Query Match: 30.30% Indels: 57
DB: 18 Gaps: 16

US-10-624-201A-2 (1-688) x US-10-424-599-133558 (1-1822)
QY 253 AsnSerSerGlyGlyGlyGluSerSerArgGlnLys-----AsnGluValAlaValGlu 271
Db 12 TCAGCTGTGGAGGAGGAGAGATAATAATAATATGATGGAGGCAACAAGAGGTGAA 71
QY 272 LeuThrAlaGlnArgGlnGluLeuGlnMetLysAlaLysLeuLeuAlaMetLeu 291
Db 72 CTCAGCACAGCACAGAGACAAGAGCTTCAGATGAAGAGTCAAAACTTGTGACCACCTGCTA 131
```

```
QY 292 GluGluValGlnGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSer 311
Db 132 GATGAGTGGAGCAAAAGGACCGACAGTATACCCACCAANTGCNAGTTGTGATAACATCA 191
QY 312 PheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAla 331
Db 192 TTTGAGCAAGCAGCGGTGTGGAGCTGCAAAAGTCTTACACAGCCCTTCGCTTAAAGACA 251
QY 332 IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer 351
Db 252 ATCTCAAGCAATTTTCGGTGTCTCAAGATGCAATCTCTTCAAAATCAAGACAACATAGC 311
QY 352 LysSerLeuGlyGluGluGluGlyLysIleGluGlySerArgLeuLysPhe 371
Db 312 AGACCTTAGGTGAAGATAATTCCTGGAGTTAGGTAGAGGGTTCGAGGCTTAGGTAT 371
QY 372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
Db 372 GTTGATCATCAACAGAGACAACAGCTGCATTA---CAGCTTGAATGATCCAACACAAC 428
QY 392 AlaTyrArgProGluArgGlyLeuProGluArgAlaValSerValLeuArgAlaTyrLeu 411
Db 429 GCTTGGAGCCCAAGAGGCTTGCTGAACGTGCTCTCTGCTCTCGTCTTGGCTTGGCTT 488
QY 412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
Db 489 TTTGAAACATTTCTTGCACCCCTATCTTAAGGACTCCGATAAGGTTATGCTTGCACAAACA 548
QY 432 ThrGlyLeuThrArgSerGlnValSerAsnThrPheIleAsnAlaArgValArgLeuTrp 451
Db 549 ACTGGGCTTACTCGGAGCCAGGTATCTAACTGGTTTATAAATGCCCGAGTTCGGCTATGG 608
QY 452 LysProMetValGluGluMetTyrLeuGluGluValLys-----AsnGlnGlu 467
Db 609 AAGCCCAATGGTAGAAGAAATGTACTTGGAGAGGTAAACAAGAGCCAAACATAGCTCC 668
QY 468 GlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaPro 487
Db 669 CAGGATACACAACAACAAGAGATCAATGAATCAAGCAAGAGGTTATGGTCAGAGGCT 728
QY 488 AsnGluGluLysHisProIleIleThrSerSerLeuLeuGluLeuAspGlyIleThrThr 507
Db 729 AATGTCTGTGCACAAAGATCAGGTGCAATGAGATTTGATCAGATTAAACATTTCTCCAATCA 788
QY 508 GlnAlaGluIleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis 527
Db 789 AAGCGAGAAAGCTTCAAGAACCAACCAACCTTCCCCAACA-----GAG 833
QY 528 HisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrVal 547
Db 834 ATCTCCAATTCCAACTCACTCCAATCTGGTTTCCATCTT-----GCT 875
QY 548 AspHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSer 567
Db 876 GACATGCAAGTCCAAACAAGCCAAAAGCACTTCTGAGATGCATCAGAAATTCCTCAGGT 935
QY 568 SerIleLeuSerSerValAspMetGluAlaLys-----578
Db 936 AGCATTCCT---TCAGTGGACATGGAATGAAGCCCTCATCATGCGGAGAAACAATAAT 992
QY 579 ---AlaArgGluSerSerAsnLysGlyPheThrAsnProLeu-----591
Db 993 ATCACTAGAGAAGGCCAGAACCAACAACAAGTTTGGCATTGAAAGCCATGCT 1052
QY 592 -----MetAlaTyrAla---MetGlyAspPheGlyArgPheAspProHis---Asp 607
Db 1053 GGGGGTTTTTGGAGCATTTTCCCAATATGGAAGACATTTGGAAGGTTTCATCATGTGCACA 1112
QY 608 GlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuPro 627
Db 1113 GAGCAACTGTCTCTAGGTTCCATGGA---AATGGTGTCTCTACTCTTGGACTTCCA 1169
QY 628 ProSer-----GluAsnLeuAlaMetPro---ValSerGlnGlnAsnTyrLeuSerAsn 644
```



```
Db 1170 CACAGTACTGAGAATAACCTTCCCTCTCAGGAACAACCAACATGATTCCTCTCACAG 1229
Qy 645 Asp-----LeuGlySerArgSerGluMetGly-----
Db 1230 AACATGCACCTGGGAATGAGAAACAGTGAATAATGATGAGTTTGTGGTGCATCAACACT 1289
Qy 654 -----SerHisTyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsn 670
Db 1290 ACCCCACCTTCTCTCACTCAGGTATCAGTTATGAGAGCATTGACATTCAAAC---AGA 1346
Qy 671 LysArgPheProThrGlnLeuLeuProAspPheVal 682
Db 1347 AAGAGGTTGTCGCGCAGTTGTTGCGGGAATTCGTG 1382

RESULT 10
US-10-425-114-7804
; Sequence 7804, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7804
; LENGTH: 1606
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700684286_FLI
US-10-425-114-7804

Alignment Scores:
Pred. No.: 3.02e-97 Length: 1606
Score: 1078.00 Matches: 259
Percent Similarity: 64.10% Conservative: 57
Best Local Similarity: 52.54% Mismatches: 102
Query Match: 30.26% Indels: 76
DB: 18 Gaps: 18

US-10-624-201A-2 (1-688) x US-10-425-114-7804 (1-1606)
Qy 253 AsnSerSerGlyGlyGlyGlyGlySerSerSerArgGlnLys---AsnGluValAlaValGlu 271
Db 3 TCGAGCTGTGGAGGAGGGGAGAAATAATAATAATATGATGGAGGCAACAACAGGAGTTGAA 62
Qy 272 LeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeu 291
Db 63 CTCAGCAGCAGACAGACAGAGCTTCAGATGAAGAAGTCAAACTGTGACCATGCTA 122
Qy 292 GluGluValGlnArgTyrArgGlnTyrHisGlnMetGlnIleValLeuSer 311
Db 123 GATGAGGTGGAGCAAAAGGTACCGACAGTATCACCAAAATGCAAGTTGTGATAACATCA 182
Qy 312 PheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAla 331
Db 183 TTTGAGCAGCAGCGGGGTGTGGAGCTGCAAGTCTTACACAGCCCTTGGCTTAAAGACA 242
Qy 332 IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer 351
Db 243 ATCTCAAAGCAATTTCTGCTGCTCAAGATGCAATCTCTTCAACAATAAAGACCACGAGC 302
Qy 352 LysSerLeuGlyGluGluGluGlyGlyGlyLysIleGluGlySerArgLeuLysPhe 371
Db 303 AAAACCTTAGGTGAAGATGATTGCTGGAGTTAAGGTAGAAGGTTTCGAGGCTTAGGTAT 362
```

```
Qy 372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
Db 363 GTTGATCATCAACTGAGACAACAACGTCATTACACAGCTTGGAAATGATCCAACAACAT 422
Qy 392 AlaTyrArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTyrPhe 411
Db 423 GCTTGGAGGCCCCAAAGAGGCTTGGCTGAACGTCGTGTTTCTGTTCTTCGAGCTGGCTT 482
Qy 412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
Db 483 TTTGAACAATTTCTTGACCCCTTATCTTAAGACTCCGATTAAGTTATGCTTGGCAACAACAA 542
Qy 432 ThrGlyLeuThrArgSerGlnValSerAsnTyrPheIleAsnAlaArgValArgLeuTyr 451
Db 543 ACTGGACTTACTCGGAGCCAGGTGTTCTAACTGTTTATAAATGCCGAGTTCGGCTATGG 602
Qy 452 LysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThr 471
Db 603 AAGCCCAATGGTAGAAGAAATGTACTTGGAGAGAGGTAAACAAAGAACCAACCAATAGCTCT 662
Qy 472 ----AsnThrSer-GlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnG1 489
Db 663 CAGGATTAACCAACAAGATCAAAAGATCAAGCAAGGAGTTATGGT-----711
Qy 489 uGluLysHisProIleIleThrSerSerLeuLeuGln---AspGlyIleThrThrThr-- 507
Db 712 -CAGAAGCTAATGCTACTGACAGGAATCTGTCGAATGAGACTTGATCATCATCAACATT 770
Qy 508 -----GlnAlaGluIleSerThrSerThrIleSerThrSerPro---ThrAlaG1 523
Db 771 CTCCAATCAAAAGCAGAAAGCTTCAACAACAACAACAACCAACCAACCACTTCCCCAACA 830
Qy 523 yAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnTh 543
Db 831 GAGATCTCTTAAT-AATTCCTTAATTCATCTCAATCAGCTGGTTTCCATCTTCTGCTGACAT 889
Qy 543 rThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLy 563
Db 890 GCAAGGCCCAACAAGCCCAATTAGAAAGCACTTCTGAAATGCAGAAC-----935
Qy 563 sPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg-----580
Db 936 ----TCTCCAGGTAGCATTCTT---TCAGTGGACATGGAATGAAGCCTCGTCATCANG 988
Qy 581 -GluSerSerAsn-----LysGlyPheThrAsnProLeuMet-----592
Db 989 AGAGACAATAATAATACTATTATTACTAGAGAGGCAACAACAACAACAAAGTTTGG 1048
Qy 593 -----AlaAlaTyrAlaMetGlyAspPh 600
Db 1049 CATTGAAAGCCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1108
Qy 600 eGlyArgPheAspProHis-----AspGlnGlnMetThrAla---AsnPh 614
Db 1109 TGAAGGTTCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1168
Qy 614 eHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProProSer---GluAsn-----631
Db 1169 TCATGGA--AATGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1225
Qy 632 -----LeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAsp-----Le 646
Db 1226 TAACCTTTTCCCTCTCAGGAACACTACTCATCAACATGGAATTCCTCTCACAGAACATGCACTT 1285
Qy 646 uGlySerArgSerGluMetGlySerHisTyrAsnArgMet-----659
Db 1286 GGGAAATGAGATCAACAACCAACCAATATGCGAATAATAGTATTTGGTGGTCCCACTCAACACTAC 1345
Qy 660 -----GlyTyrGluAsnIleAspPheGlnSerGlyAs 670
Db 1346 TCCACCTTCTAATTCCTCACTCAGCACCAGCTTATGAGCATTTGACATTTATTTCAAAACAG 1405
```



```
Db 2331 GAGGGAGCAAGCCCGGTGGAGAGAGCGCGCGCGCGTCAATGGCGGTGTGTT 2390
Qy 481 -----GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThr 496
Db 2391 GACAGTCCGCGAGATGGACAGCAGCGCGGCACATGGAGAGCGCGCGGTATCAT 2450
Qy 497 SerSerLeuLeuGln-AspGlyIleThrThrThrGlnAlaGluIleSerThrIle 516
Db 2451 CCGTCGTCTCGAGTCGCGCGGTGACACAGCGCGAGCGCGGTCTACAGCAGCAC 2510
Qy 516 eSerThrSerProThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeu-- 535
Db 2511 GACGAGGATGCGGC-GCCGCGCGCGCTCCAGCAGAGAGCTGAAGAAGCGGACGGA 2569
Qy 536 -----GlySerPheAsnMetAspAsnThrThrThrValAspHisIleG1 551
Db 2570 GGACGACGACGAGCGGGTTCAGGTGTCGAGTGGCCACGCTGCACGCGCATGCGC 2629
Qy 551 uAsnAsnAlaLysLysGlnArgAsnAsp-----MethHisLysPheSe 565
Db 2630 GGCGCGCGCGCGAGCAGCAGCAGAGGTGAGCCACCGGAGCTCCTCATGAAGTTTCA 2689
Qy 565 rProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLy 585
Db 2690 GGAGAGCGGC---AGCGCGCGCGCGCGCGCGCGCGCGCGCGGACACACACGCA 2746
Qy 585 sGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspPr 605
Db 2747 ACACCATGGCGCGTTCGCTACTCGTGTTCGCGCGCGCGCGCGCGCGCGCGTTC- 2801
Qy 605 oHisAspGlnGlnMetThrAlaAsnPheHis-----GlyAsnAsnGlyValSe 621
Db 2802 -GCCACGCGCAGTTCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTG 2860
Qy 621 rLeuThrLeuGlyLeuPro---ProSerGluAsnLeuAlaMetProValSerGlnGlnAs 640
Db 2861 GCTCAGCGTCCGCTCCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG-TC 2902
Qy 640 nTyrLeuSerAsnAspLeuGlySerArgSerGluMetGlySerHisTyrAsnArgMetG1 660
Db 2903 GTTCCTCATGACCCAGC---AGCAAGCGCAGCAGCGCGCGCGCGCGCGCGCGCG 2959
Qy 660 yTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAs 680
Db 2960 CGGTACGACATGAACATGCAGAGC---ACAAAGTCTCTTCGCGGTCTCAGCTCATGAGA 3016
Qy 680 pPheVal 682
Db 3017 CTTCGTG 3023
```

RESULT 12

```
US-10-424-599-63874
; Sequence 63874, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 63874
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28692C.1
US-10-424-599-63874
```

Alignment Scores:

```
Pred. No.: 4.89e-93 Length: 1892
Score: 1037.00 Matches: 254
Percent Similarity: 62.88% Conservative: 56
Best Local Similarity: 51.52% Mismatches: 108
Query Match: 29.10% Indels: 76
DB: 18 Gaps: 18
```

US-10-624-201A-2 (1-688) x US-10-424-599-63874 (1-1892)

```
Qy 253 AsnSerSerGlyGlyGlyGlyGlySerSerArgGlnLys---AsnGluValAlaValGlu 271
Db 25 TCGAGCTGTGGAGAGCGGAGAAATAATAATAATGATGGAGGCAACAAAGAGTGTGA 84
Qy 272 LeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeu 291
Db 85 CTGACGACAGCAGACAGACAGAGCTTCAGATGAAGAGTCAAAACTTGTGACCATGCTA 144
Qy 292 GluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSer 311
Db 145 GATGAGGTGGAGCAAGAGTACCGACAGTATCACCACCAATGCAAGTTGTGATAACATCA 204
Qy 312 PheGlnGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAla 331
Db 205 TTTGAGCAAGCAGCGGGTGTGGAGCTGCAAAAGTCTTACACAGCCCTTGCCTTAAAGACA 264
Qy 332 IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer 351
Db 265 ATCTCAAGCAANTTTCGCTGTCTCAAGATGCAATCTCTTCACAAATAAAGACACGAGC 324
Qy 352 LysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLysPhe 371
Db 325 AAAACCTTAGTGAAGATGATTGCTGGAGATTAAAGTAGAAGGTTTCGAGGCTTAGGTAT 384
Qy 372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnIleGlyMetMetGlnProAsn 391
Db 385 GTTGATCATCACTGACACCAACACGTCATTAACACAGCTTGAATGCGGACGCGTGGG 444
Qy 392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
Db 445 CGNACGCGTGGGCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 504
Qy 412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
Db 505 TTTGACACATTCTTGCACCCCTTATCTTAAGACCTCCGATAGGTTATGCTTGTCAAAACAA 564
Qy 432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
Db 565 ACTGACCTTACTCGAGCCAGGTGCTAACTGGTTTATAATGCCCGAGTTTCGGCTATGG 624
Qy 452 LysProMetValGluGluMetTyrLeuGluValLysAsnGlnGluAsnSerThr 471
Db 625 AAGCAATGTTAGAGAAATGTTACTTGGAGAGAGTAAACAGAACCAACAAATAGCTCT 684
Qy 472 -----AsnThrSer-GlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnG1 489
Db 685 CAGGATAACACCAACAAAAGATCAAAAGATCAAGCAAGAGTATATGCT----- 733
Qy 489 uGluLysHisProIleIleThrSerSerLeuLeuGln---AspGlyIleThrThrThr-- 507
Db 734 -CAGAAGCTAATGCTACTGCACAGCAATCTGGTCAATGAGACTTGATCATCAACATT 792
Qy 508 -----GlnAlaGluIleSerThrSerIleSerThrSerPro---ThrAlaG1 523
Db 793 CTCCAATCAAAAGCAGAAAGCTTCAACACCAACCAACCAACCAACCACTTCCCCAACA 852
Qy 523 yAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnTh 543
Db 853 GAGATCTCTAAT-AATTCCCATTAATTCACCTCAATCAGCTGGTTTCCATCTTGTGACAT 911
Qy 543 rThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMethHisLy 563
Db
```



```
Db 1599 GCGCGAGCCGGCGCTGGCGGAGACCGCGGACCGCGCGCTCGCGGGCGCGCACCGCTG 1658
Qy 366 GlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIle 385
Db 1659 GGGTCCCGCTCCGCTACATCGACACACAGCTCCGGGAGAGCGCGCTCGACGAGCTC 1718
Qy 386 GlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSer 405
Db 1719 GGCATGATGACAGGCGCGCGCTGGCGGCCCGCGCGCTCCCGAGCGCTCCGCTCTCC 1778
Qy 406 ValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLys 425
Db 1779 ATCTCCGCGCTGGCTTTCGAGACATTCCTGACCCATCCCAAGGATTCGGACAAG 1838
Qy 426 IleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsn 445
Db 1839 ATCATGCTCGCAGCAACCGGGCTCACAGGAGTCAAGTGTCCATTGGTTTCATCAT 1898
Qy 446 AlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsn 465
Db 1899 GCGAGGGTGGCGCTGTGGAAGCTATGTTGGAGAGATGTACCTGGAGAGACAAAGGAC 1958
Qy 466 GlnGluGln-----AsnSerThrAsnThrSerGlyAspAsnLysAsnLysGlu 481
Db 1959 CAGACGCTGGAGCGCGCGGACGAGCGGCAAGTCCGGTGGCGCGGCGCAAGAGC 2018
Qy 482 ThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuGln 501
Db 2019 AGCAGACCGTCGACGGGCTCATCCGAGGCGGACGATGTCAAGTCCCGCGGTGGC 2078
Qy 502 AspGly-----IleThrThr-ThrGlnAlaGluIle----- 511
Db 2079 GTGGAGGCGCGCGCGCGCGGAGAGCGGTCACCAACAAGGCGATCCATGGCTCC 2138
Qy 512 -----SerThrSerThrIleSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHi 527
Db 2139 TCCTGTCTCGAGCTCGCGCGCGGCGATCACCAGCAGTCCACGCGGGTCTTACGAGC-- 2196
Qy 527 sHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThr 546
Db 2197 -----ACGNTGAAGACGCGCGGACGAGC 2220

RESULT 14
US-10-425-115-97705
; Sequence 97705, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION: Thomas J.
; APPLICANT: La Rosa, David K.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 97705
; LENGTH: 2970
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_20618C.1
US-10-425-115-97705

Alignment Scores:
Pred. No.: 6,11e-89 Length: 2970
Score: 999.00 Matches: 246
Percent Similarity: 56.21% Conservative: 80
Best Local Similarity: 42.41% Mismatches: 155
Query Match: 28.04% Indels: 99
DB: 20 Gaps: 18
```

```
US-10-624-201A-2 (1-688) x US-10-425-115-97705 (1-2970)
Qy 2 TyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHisAsn 21
Db 1034 TACTTACACAGGTGGCGCC---GGCACGAGATCCAGGCC----- 1069
Qy 22 HisGlyAsnSerAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsnTyr 41
Db 1070 -----AGCACCGACGGCTCGACAGCGCTGTACCTCATGAACCCG---AGCTAC 1114
Qy 42 MetGlnGlyTyrThrThr-----SerAspThrGlnGlnGlnGlnLeuLeuPhe 58
Db 1115 GCGCGCGCTACGCTGACGACGCGCGCTCCACGCGCGGGCGGCAACATGATGCTC 1174
Qy 59 LeuAsnSerSerProAlaAla-----SerAsnAlaLeuCysHisAlaAsnIleGln 75
Db 1175 CTCAACTCGGCGGTGACCATCATGACCGCGGCTCTTGGCGACACCAACGACGAG 1234
Qy 76 HisAlaProLeuGlnGlnHisPheValGlyValProLeuProAlaValSerLeuHis 95
Db 1235 TCGCGCGCGCGCGGACGACACTTCTCGGCATCCCGCTCCAGGTC----- 1282
Qy 96 AspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnGlnAsnGlnSerGln 115
Db 1282 ----- 1282
Qy 116 GlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyIleThrThr 135
Db 1283 CAGCGCGCGCGCGCTCGGCTACAACCTGTGGACCCCGGCATGGCTGGCGGCGCGCG 1342
Qy 136 AspLeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGln 155
Db 1343 GACATGTCCGCGCAGCGCGGCGGCGCAGCTCCC-----GGC 1378
Qy 156 GlnGlnGlnGlnGlyLeuSer-----LeuSerLeuSerProGlnLeuGlnGlnGln 173
Db 1379 GCGCGCGGTGCTGCTGGGTGAGCGCGCTGCTACGCTGTCTCCCGCGCGCGCGCGCG 1438
Qy 174 IleSerPheAsnAsnIleSerSerSerSerProArgThrAsnAsnValThrIleArg 193
Db 1439 GTCACC-----GTGGCGCGGTGGCGCGCTGCCCTGCCACGAGGGGCAAGTACCTG 1486
Qy 194 GlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLysThrLysLeuAlaAla 213
Db 1487 GCGGTGTCCGCCACCGCGCAGGGGCGAGATGTCATGAGTCCAAGTACCTCAAGCGCG 1546
Qy 214 GlnGlnLeuLeuAspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGln 233
Db 1547 CAGGAGCTGCTCGACGAGGTGGTCAGC---GTCAGCAAGGGCGCTC-----GAGGACGCT 1597
Qy 234 LysLysAspAsnSerMetAsnLysGluSerMetProLeuAla-----SerAspValAsn 251
Db 1598 AAGAGCGCGCTGCCCGCGCGACCAAGAGCTGGCCCGGTGAAGAAGAGGAGGACTCG 1657
Qy 252 ThrAsnSerSerGlyGlyGlyGlu-----SerSerSerArgGlnLysAsnGluVal 268
Db 1658 GAGGCGCGCTCAGGTGGCGCGCACCGACGACGCTGCCCGCGCAAGAGCGGGGTGTC 1717
Qy 269 AlaValGluLeuThrThrAlaGlnArgGlnGlnLeuMetLysLysAlaLysLeuLeu 288
Db 1718 GCGCAGAGATGTCCACGCGGCGGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1777
Qy 289 AlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIle 308
Db 1778 AACATGCTTACGAGGTGGAGACGCGGTACCGGACGTACCGGCGGACATCAGCGCGGTG 1837
Qy 309 ValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAla 328
Db 1838 TCGTGTGCTTCGAGGCGCGCGCGCGCGCGCGGTGCGGCGGAGTACACGCGCTGGCG 1897
Qy 329 LeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLys 348
Db |||||
```

```
Db 1898 CTGGCACCATTCTCGCGCAGTTCCTGGGTGTCTGGGACGCGATCGCGTCGCGAGTGC 1957
Qy 349 AlaThrSerLysSerLeuGluGluGluGly-----LeuGlyGlyLysIleGlu 365
Db 1958 GCAGCGAGCGCGCGCTGGCGGAGGACGCGAGCGCGCCCTCGCGCGGCGCGCAGCGTG 2017
Qy 366 GlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnIle 385
Db 2018 GGGTCCCGCTTCGCTACATCGACCACAGCTCGGCGAGGAGCGCGCTGCGAGCGTCTC 2077
Qy 386 GlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSer 405
Db 2078 GGCATGATGACGAGCGCGCGCTCGCGCGCCAGCGCGCTCCCGGAGCGCTCGCTCC 2137
Qy 406 ValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLys 425
Db 2138 ATCTTCGCGCGCTGGCTTCGAGCACTTCCTGCACCCATACCCCAAGGATTCGGACAAG 2197
Qy 426 IleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsn 445
Db 2198 ATCATGCTCCCAAGCAACCGGGCTCACCAGGAGTCAGGTGCCAATTCGTTCAATCAAT 2257
Qy 446 AlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsn 465
Db 2258 GCGAGGTCGCGCTGTGGAAGCCTATGCTGAGGAGATGTACCTGGAGGAGACAAAGGAC 2317
Qy 466 GlnGluGln-----AsnSerThrAsnThrSerGlyAspAsnLysAsnLysGlu 481
Db 2318 CAGGACGCTCGAGCGCGCAACGACGAGCGCAAGTCCCGTGGCGCGCGCAGCAAGAGC 2377
Qy 482 ThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerLeuLeuGln 501
Db 2378 AGCGACCGCTCGACGCGCTCATGCCGAGGCGGAGCGGTGATGCCAAGTCCGGGTGCGC 2437
Qy 502 AspGly-----IleThrThr-ThrGlnAlaGluLe----- 511
Db 2438 GTGGGCGCGCGCGCGCGGCGGAGAGCGCGTCCACCAAGGCGCATCCATGGCTCC 2497
Qy 512 -----SerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHi 527
Db 2498 TCCCTGTCGAGCTCGCGCGCGGCGATCACAGCAGTCCACCGCGGGTCTCAGACG-- 2555
Qy 527 sHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThr 546
Db 2556 -----ACGATGAAGACGCGCGAGCG 2579

RESULT 15
US-10-425-114-9882
; Sequence 9882, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9882
; LENGTH: 1716
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700875318_FLI
US-10-425-114-9882
Alignment Scores:
, Pred. No.: 8,51e-88 Length: 1716
```

```
Score: 984.00 Matches: 230
Percent Similarity: 63.72% Conservative: 44
Best Local Similarity: 53.49% Mismatches: 90
Query Match: 27.62% Indels: 66
DB: 18 Gaps: 14

US-10-624-201A-2 (1-688) x US-10-425-114-9882 (1-1716)
Qy 306 GlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThr 325
Db 2 CAAATTTGGTATCATCATTTGAGCAAGCAGCGGGTTATGGGCGGCAAAATCTTACACT 61
Qy 326 GlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGlu 345
Db 62 GCCCTTGCCTTAAATACTATCTCAAGCAATTCAGGTGCTGAAAGATGCAATCTCTGCA 121
Qy 346 GlnValIleAlaThrSerLysSerLeuGlyGluGluGlyLeuGlyGlyLysIleGlu 365
Db 122 CAAATCAAAGCCACGAGCAAGACATTTGGGTGAAGATGATTCCTTAGGAGTTAAAGTAGAA 181
Qy 366 GlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnIle 385
Db 182 GGGTCAAGGCTTAGGTATGTTCACCACCATCTTCGGCAACAAGGGCAGCTTCAGCAACTA 241
Qy 386 GlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSer 405
Db 242 GGAATGATTCAGCCCAATGCATGAGGCGCCAGAGAGCTTGCCTGAACGAGCTGTTTCC 301
Qy 406 ValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLys 425
Db 302 ATTCTTCGGGCTTGGCTTTTGAGCATTTCTCTCACCCATATCCAAAGGAGCTCCGATAAA 361
Qy 426 IleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsn 445
Db 362 GTTATGCTTGTCTAAACAAACCTGACATTTCTAGAGCGCAGGTGTCAACATCGTTTATCAAT 421
Qy 446 AlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsn 465
Db 422 GCCCGAGTTCGGCTTTGGAAGCCCAATGGTTGAAGAAATGTACTTGAAGAAATCAAGGAA 481
Qy 466 GlnGluGlnAsn-----SerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsn 483
Db 482 CATGAACAAGTAAATGCCTCAGAAAAACACCAAAATCAAGGAATCGAGCAAGAGTGGGT 541
Qy 484 IleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeuGlnAsp-- 502
Db 542 TCAACAGCAAAATGTTGCACCCGAATCAGGTGCCATTAAACTCGATCATCTCTCAATCAAG 601
Qy 503 -----GlyIleThrThrGlnAlaGluIleSerThrSerThrIleSer 517
Db 602 CAAGAAAGCTTCAACACCAACAAACACTTCTCCAACCGAGATCTCAACAAACTCT--TCA 658
Qy 518 ThrSerProThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySer 537
Db 659 ATGCTCTCCA---ATGGGAGGGTCCCTCAATCCCCCTCTGTTTCCATCTTTCGAGGATCA 715
Qy 538 PheAsnMetAspAsnThrThrThrValAspHisIleGluAsnAsnAlaLysLysGln 557
Db 716 TCTGACATGCAA-----ATTAGGAGTCCAAAGAAACCA 748
Qy 558 ArgAsnAspMetHisLysPheSerProSerSerIleLeuSerValAspMetGlu--- 576
Db 749 AGAAGTTCGGAATTCAAAACTCTCCAAGTAGTATATCTC--TCAGTGGGCATGGAAATG 805
Qy 577 -----AlaLysAlaArgGluSerSerAsnLysGlyPheThr 588
Db 806 AAGCATAATAATAATGGTGTGATCATGCAATAATAATAGAGAGCCACACACAAAGTTTGGC 865
Qy 589 Asn-----ProLeuMet----- 592
Db 866 ATGAAAGGCCACCACCAAGAGATGATACCTTTTGTATGACTAGCAATGCAAAACCATGGT 925
Qy 593 -----AlaAlaTyrAlaMetGlyAspPheGly---ArgPheAspProHisAspGln 608
```

```
Db      926  GGAGGGTTTCGAGCGTTCCACATGGAGACATGGAGCAGATTTAATGTCACTACTGAG 985
Qy      609  GlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProPro 628
Db      986  CAACTGGCTTCAAGGTTCCATGGA--AATGGTGTTCCTCTCACACTTGGACTTCCCCAC 1042
Qy      629  SerGluAsnLeuAlaMetProValSerGlnGlnAsn---TyrLeuSerAsnAsp----- 645
Db      1043  AATGAGAACCTTTCTATGCCGGAACCTCAACAACATGGATTCTCTCACAGAATATACAC 1102
Qy      646  LeuGlySerArgSerGluMetGlySerHisTyrAsnArgMet----- 659
Db      1103  TTGGGAAGGAGACTTGAATGGGAACAATAATGGAAATGAGTTTGTGCTATCAACACTCCA 1162
Qy      660  -----GlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArg 672
Db      1163  CCTTCTTCTCACTCAGGAACCACTTATGAGAGCATTTGACATTCAAAACAGAAAGAGGTTT 1222
Qy      673  PheProThrGlnLeuLeuProAspPheVal 682
Db      1223  GTTGCTCATCAGTTATTGCCAGATTTCGTG 1252
```

Search completed: July 29, 2005, 19:19:10
Job time : 1191 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 14:13:12 ; Search time 5760 Seconds
(without alignments)
4546.566 Million cell updates/sec

Title: US-10-624-201a-2

Perfect score: 3563

Sequence: 1 MYVQGTSDNTNIQADHQQRH.....GNKRFPQTLLPDPFTGNLGT 688

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delep 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10624201/runat_27072005_122958_10917/app_query.fasta_1.839
-DB=EST -QFMT=fastcap -PREFIX=rrst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US10624201.OCGN 1.1 4352 @runat_27072005_122958_10917 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hrc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1149	32.2	685	2 AW034617	AW034617 EST278301
2	1102	30.9	756	4 BG593861	BG593861 EST492539
3	1057	29.7	964	2 CK262102	CK262102 EST708180
4	1043.5	29.3	644	2 AW930221	AW930221 EST340678
5	1003.5	28.2	630	4 BG129304	BG129304 EST474950
6	981.5	27.5	610	7 CO909070	CO909070 BJ02039F0
7	966	27.1	699	5 BQ119770	BQ119770 EST605346
8	963.5	27.0	1965	9 CL978605	CL978605 OSIFCC032
9	961	27.0	708	4 B1920940	B1920940 EST540875

10	930	26.1	623	2 AW441343	AW441343 EST310739
11	920	25.8	553	1 AI773253	AI773253 EST254353
12	898	25.2	827	1 CV241175	CV241175 WS02511.B
13	895	25.1	789	4 B1921008	B1921008 EST540943
14	892.5	25.0	850	6 CB292855	CB292855 UCRCS01.0
15	889.5	25.0	752	4 BM555737	BM555737 EST588759
16	884.5	24.8	727	4 B1176426	B1176426 EST517371
17	875	24.6	783	2 AW688195	AW688195 NF004E06S
18	874	24.5	669	5 BQ148285	BQ148285 NF065F11F
19	871	24.4	740	7 CO111273	CO111273 GR_EB004
20	870	24.4	568	2 AW441354	AW441354 EST310750
21	866.5	24.3	844	7 CO101205	CO101205 GR_EB002
22	864	24.2	667	1 AI485358	AI485358 EST243679
23	839.5	23.6	1803	9 CL961014	CL961014 OSIFCC005
24	839	23.5	554	1 AI897675	AI897675 EST267118
25	835	23.4	781	7 CO098143	CO098143 GR_Ea22F
26	834	23.4	1064	7 CV477853	CV477853 57758.1.D
27	833	23.4	495	2 BE460264	BE460264 EST411755
28	829.5	23.3	777	6 CB343619	CB343619 CA32EN000
29	829.5	23.3	1954	3 CNS0ADW7	BN813490 Arabidops
30	823	23.1	523	1 AI898911	AI898911 EST268354
31	823	23.1	530	1 AI487525	AI487525 EST245847
32	823	23.1	645	4 BG887850	BG887850 EST513701
33	816	22.9	480	5 BQ509372	BQ509372 EST616787
34	801	22.5	636	4 BG887971	BG887971 EST513822
35	799.5	22.4	889	7 CO366367	CO366367 RTK1_27.A
36	795	22.3	576	1 AI894617	AI894617 EST264060
37	793.5	22.3	636	2 AW442048	AW442048 EST311444
38	789	22.1	537	5 BM954132	BM954132 ga070812.
39	781	21.9	572	5 BP034454	BP034454 BP034454.
40	779	21.9	1278	9 CL980113	CL980113 OSIFCC034
41	778	21.8	645	5 BU043836	BU043836 PP_LEA001
42	777	21.8	555	4 BM142730	BM142730 sa354b05.
43	775.5	21.8	650	4 BG457258	BG457258 NF101F03P
44	774	21.7	616	6 CA839352	CA839352 MCT026810
45	768.5	21.6	747	6 CB671354	CB671354 OSUNE040

ALIGNMENTS

RESULT 1
AW034617
LOCUS EST278301 tomato callus, TAMU Lycopersicon esculentum cDNA clone .
DEFINITION CLEC33H16, mRNA sequence.
ACCESSION AW034617
VERSION AW034617.1 GI:5893373
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 685)
REFERENCE 1 Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato callus tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES

source
1..685
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEC33H16"

```

/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRP"
/clone_lib="tomato callus, TAMU"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

```

ORIGIN

Alignment Scores:
 Pred. No.: 1.14e-107 Length: 685
 Score: 1149.00 Matches: 222
 Percent Similarity: 98.70% Conservative: 5
 Best Local Similarity: 96.52% Mismatches: 1
 Query Match: 32.25% Indels: 2
 DB: 2 Gaps: 1

US-10-624-201A-2 (1-688) x AW034617 (1-685)

```

QY 321 AlaLysSerThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLys 340
DB 1 GCAAAATCATACACTCAATTAGCTTTGCGATGCAATTCAAAGCAATTCAGATCGCTAAAG 60
QY 341 AspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGlyLeu 360
DB 61 GATGCAATTAGTGAGCAAGTGAAGGCGACGAGCAAGATTTAGGTGAAGATGAAGGCTTG 120
QY 361 GlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnArg 380
DB 121 GGAGGGAAATTCAGGCTCAAGACTCAATTTGTGGACCATCATCTAGGCAACACGC 180
QY 381 AlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuPro 400
DB 181 GCGCTGCAACAGCTAGGAATGATGCAACCAAAATGCTTTGGAGACCCCAAGAGGTTTACCT 240
QY 401 GluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrPro 420
DB 241 GAAAGAGCTGCTCTGCTTCGTGCTTGGCTTTTCGAGCATTTCTTCATCTTACCCC 300
QY 421 LysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSer 440
DB 301 AAAGACTCAGACAAATCATGCTTGCTAAGCAAAACGGGCTAACAGGAGCCAGGCTCTCT 360
QY 441 AsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeu 460
DB 361 AACTGGTTCAATTAATGCTCGAGTTCGATTATGGAAGCCAAATGGTTGAAGAAATGTACTTG 420
QY 461 GluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
DB 421 GAAGAAGTGAAGAAATCAAGAACAAACACAGTAGTAGTAATCTTCAGAGATACAAACAA 480
QY 481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
DB 481 GAGACGAATATAAGTGCTCCAAATGAAGAGAAACCAATATTATTACTAGCACCTTATTA 540
QY 501 GluAspGlyIleThrThrGlnAlaGluIleSerThrSerThrIleSerThrPro 520
DB 541 CAAGATGGT-----ACTACTCAAGCAGAGAAATTTCTACCTCAACTATTTCACCTCCCT 594
QY 521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
DB 595 ACTGCAAGGCTTCACCTTCATCATGCTCACAATTTCTCTTCCTTCTGTTCAATTCACATG 654
QY 541 AspAsnThrThrThrValAspHisIle 550
DB 655 GAGAATACTACTACTACTGTTGATCATATT 684

```

RESULT 2

BG593861

LOCUS

756 bp mRNA linear EST 07-MAR-2003

DEFINITION

EST492539 cSTS Solanum tuberosum cDNA clone cSTS5p9 5' sequence, mRNA sequence.

ACCESSION BG593861 GI:13612001

VERSION EST.

SOURCE Solanum tuberosum (potato)

ORGANISM

Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; famids; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 756)

REFERENCE

AUTHORS

van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chieningo,A.,

Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.

Generations of ESTs from sprouting potato eyes

Unpublished (2000)

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

This clone can be obtained from the University of Arizona Genomics

Institute. Orders can be made through URL:

http://genome.arizona.edu/orders/

Seq primer: MJ3F-R.

FEATURES

source

```

1..756
    Location/Qualifiers
        /organism="Solanum tuberosum"
        /mol_type="mRNA"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="cSTS5p9"
        /tissue_type="sprouting eyes from tubers"
        /dev_stage="12-14 weeks post harvest"
        /lab_host="SOLs"
        /clone_lib="cSTS"
        /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."

```

ORIGIN

Alignment Scores:

Pred. No.: 9.41e-103 Length: 756
 Score: 1102.00 Matches: 210
 Percent Similarity: 99.06% Conservative: 0
 Best Local Similarity: 99.06% Mismatches: 2
 Query Match: 30.93% Indels: 0
 DB: 4 Gaps: 0

US-10-624-201A-2 (1-688) x BG593861 (1-756)

QY 1 MetTyrTyTGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnArgHis 20

DB 119 ATGTACTATCAAGGAACCTCGGATATATACTAATATATACTCAAGTCATCATCAACACATCAT 178

QY 21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40

DB 179 AATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTATTGATGAACCTTACCAAT 238

QY 41 TyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGlnLeuLeuPheLeuAsn 60

DB 239 TATATCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCTTACTTTTCTCTGAAT 298

QY 61 SerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGln 80

DB 299 TCTTCCAGCAGCAAGCAACGCGCTTTCGCAATATACACACGCGCGCTGCAA 358

QY 81 GlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsnHis 100

DB 359 CAGCAGCATTTGTGCGTGTGCTCTTCGCGCAGTAGTATGTTGCACCATCAGATCATCAT 418

QY 101 HisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValPro 120

```
Db 419 CATGGACCTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCA 478
Qy 121 SerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGln 140
Db 479 TCGTCGAGCGGGGTTTCTGCCAGCTCAATGTGGCGGATCACACGGACTTGGCGTCTCAA 538
Qy 141 LeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGly 160
Db 539 TTGGCGTTTCAGAGCGCGATTCCGACACCAACACACCGACGACGACCAACACAGCAGGC 598
Qy 161 GlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 599 GGTCTATCTCTTAAGCCCTTCTCCCTACGCTACCAACAGACAAATAGTTTCAATAACAATATT 658
Qy 181 SerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySerSer 200
Db 659 TCATCCTCATCACCAAGACAAATATGTTACTATTAGGGGACATTTAGATGGGAAGTCT 718
Qy 201 SerAsnMetValLeuGlySerLysLeuLysAla 212
Db 719 AGCAACATGGTTTATAGGCTCTAAGTATCTGAAAGCT 754
```

RESULT 3

```
CK262102 964 bp mRNA linear EST 03-AUG-2004
LOCUS EST708180 potato abiotic stress cdna library Solanum tuberosum cDNA
DEFINITION clone POABE14 5' end, mRNA sequence.
```

```
ACCESSION CK262102
VERSION CK262102.1 GI:39819080
KEYWORDS EST.
```

```
SOURCE Solanum tuberosum (potato)
```

ORGANISM

```
Solanum tuberosum
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
```

REFERENCE

```
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)
```

```
COMMENT Other ESTs: EST708181
```

```
Contact: Robin Buell
```

```
The Institute for Genomic Research
```

```
9712 Medical Center Dr, Rockville, MD 20850, USA
```

```
Email: potato-array@tigr.org
```

```
Clones can be requested from the University of Arizona Genomics
```

```
Institute via http://genome.arizona.edu/orders/.
```

```
Seq primer: ATT TAG GTG ACA CTA TAG.
```

```
Location/Qualifiers
```

FEATURES

```
source
```

```
1. .964
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POABE14"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cdna library"
/note="vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
```

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

```
Alignment Scores: 6.08e-98 Length: 964
Pred. No.: 1057.00 Matches: 226
Score: 77.35% Conservative: 37
Percent Similarity: 66.47% Mismatches: 51
Best Local Similarity: 29.67% Indels: 27
Query Match: 9
DB: 7
Gaps: 9

US-10-624-201A-2 (1-688) x CK262102 (1-964)

Qy 237 AsnSerMetAsnLysGluSerMetProLeuAlaSerAspValaAsnThrAsnSerSerGly 256
Db 8 CACAATTTGGACAAATGAATTAATATCTTTGGCTAGTGAT----- 46

Qy 257 GlyGlyGluSerSerArgGlnLysAsnGluValaValaGluLeuThrThrAlaGln 276
Db 47 ---GATGTTGAAAGTAGCAGCCAAAAAAT---ATTGTTGTTGAATCTTACTACAGCTCAA 100

Qy 277 ArgGlnGluLeuGlnMetLysLysAlaLeuLeuLeuAlaMetLeuGluGluValGluGln 296
Db 101 AGACAAGAACTTCAAAATGAAGAAGCCAAAGCTTTGTAGCATGCTTGTAGAGTGGATCAA 160

Qy 297 ArgTyArgGlnTyHisHisGlnMetGlnIleLeuValLeuSerPheGluGlnVala 316
Db 161 AGTATAGACAATACCATCACCAATGCAATGATTTGCCAATCATTTTGAGCAACAACA 220

Qy 317 GlyIleGlySerAlaLysSerTyThrGlnLeuAlaLeuHisAlaIleSerLysGlnPhe 336
Db 221 GGAATTTGGATCATCAAAATCATACACAACTTGTCTTGACACACAATTTTGAAGCAATTT 280

Qy 337 ArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu 356
Db 281 AGATGTTTAAAGATGCAATTTCTGGCAATTAAGGACACAAAGTAAATCTTTAGGGGAA 340

Qy 357 GluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisLeu 376
Db 341 GAAGAGAACATTGGAGGCAAAATTTGAAGGATCAAGTTGAAATTTGTGGATCATCATTTA 400

Qy 377 ArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaThrArgProGln 396
Db 401 CGCCAAACACGTCACCTACCAACAATTAGGGATGATGCAAAACCAATGCGATGGAGGCCACAG 460

Qy 397 ArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeu 416
Db 461 AGAGGTTTGGCCGAAGAGCGGTTTCGGTTCTCCGCGCTTGGCTTTTCGAGCATTTCTT 520

Qy 417 HisProTyProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArg 436
Db 521 CATCCGTATCCCAAGATTTCAGATAAAATCATGCTTGTCTAAGCAACAGGCTTAACAAGG 580

Qy 437 SerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGlu 456
Db 581 AGCCAGGTATCAAAATTTGGTTTATAAAATGCTAGAGTTAGACTATGGAAGCAATGTTAGAA 640

Qy 457 GluMetTyLeuGluGluValLys-----AsnGlnGlnGlnAsnSerThrAsnThrSer 474
Db 641 GAAATGTACATGGAAGAGTGAAGAAACCAATCAAGACAAATAATTTGAGCCCTTAATAAC 700

Qy 475 GlyAsnLysAsnLysGluThrAsnIleSerAlaProAsnGlnGluLysHisProIle 494
Db 701 AATGAATTTGTTGGTTCAAAATCAAGTTTCCA-----CAAGAGAAATTAACCA--- 748

Qy 495 IleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSer 514
```

```
Db 749 ATTAGTAGCAATATTATTCATAAT-----GCTTCTCCAAATGATATTCTACTTCC 799
Qy 515 ThrIleSerThrSerProThrAlaGlyAlaSerLeuHisHisAlaHis-----Asn 531
Db 800 ACCATTTCACATCTCCGACGCGCGGTCGATTCGCGCTCAGACGGTTCAGGT 859
Qy 532 PheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIleGlu 551
Db 860 TTCTCTTCATTAGTTCATTAAACATGGAGAAC-----ATTGATGATCAAGG 907
Qy 552 AsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerGerrileLeuSer 571
Db 908 AACAAAC--AAAAAGGCAAGAAATGAGATGCA-AAATTGTTCTACTAGTACTATTCTCTCA 963

RESULT 4
AW930221
LOCUS EST340678 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF41L1 5', mRNA sequence.
ACCESSION AW930221 GI:8105622
VERSION 1
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 644)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.B., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
source
Location/Qualifiers
1..644
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF41L1"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/clone_lib="tomato fruit mature green, TAMU"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
ORIGIN
Alignment Scores:
Pred. No.: 7,95e-97 Length: 644
Score: 1043.50 Matches: 203
Percent Similarity: 96.76% Conservative: 6
Best Local Similarity: 93.98% Mismatches: 4
Query Match: 29.29% Indels: 3
DB: 2 Gaps: 2
US-10-624-201A-2 (1-688) x AW930221 (1-644)

Qy 362 GlyIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnArgAla 381
Db 3 GGGAAAATTGAGGCTCAAGACTCAAAATTTGTGGACCATCATCTAAAGGCAACACGCGC 62
```

```
Qy 382 LeuGlnGlnIleGlyMetMetGlnProAsnAlaTtpAtgProGlnArgGlyLeuProGlu 401
Db 63 CTGCACACACTAGGAATGATGACCAAAATGCTTGGAGACCCCAAGAGAGGTTTACTCTGAA 122
Qy 402 ArgAlaValSerValLeuArgAlaTtpLeuPheGluHisPheLeuHisProTyrProLys 421
Db 123 AGAGCTGTCTCTGTCCTTGGCTTTTCGAGCAATTTCTTCATCCTTACCACAAA 182
Qy 422 AspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsn 441
Db 183 GACTCTAGACAAAATCATGCTTCTAAGCAAAACGGGGCTAAACAGGAGCCAGGTCCTTAAC 242
Qy 442 TtpPheIleAsnAlaArgValArgLeuTtpLysProMetValGluGluMetTyrLeuGlu 461
Db 243 TGGTTCAATTAAGTCTCAGTTTCATATGAGAGCCAAATGGTTGAAGAAATGACTTTGGAA 302
Qy 462 GluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGlu 481
Db 303 GAAGTGAAGAATCAAGAACAAAACAGTAGTAATACTTTCAGGAGATAACAAAACAAAGAG 362
Qy 482 ThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeuGln 501
Db 363 ACGAATATAAGTCTCCAAATGAAGAGAGAAACCAATATTATTACTAGCAGCTTATTACAA 422
Qy 502 AspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerProThr 521
Db 423 GAAGGT-----ACTACTCAGCAGAAATTTCTACCTCACTATTTCACATCCCTACT 476
Qy 522 AlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAsp 541
Db 477 GCAGGTGCTTCACTTCATCATGCTCACAACTTCTCCTCTCTGTTCAITCAACATGGAG 536
Qy 542 AsnThrThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsn---App 560
Db 537 AATACTACTACTACTGTTGATCATATTTGAAACAAACGCGAAAAGCCAAAGAAATCATGAC 596
Qy 561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGlu 576
Db 597 ATGCACAAAGTTTCTTCAAGTAGTAATATCTTTTATCTGTTGAATGAA 644
```

RESULT 5
BG129304
LOCUS EST474950 tomato shoot/meristem Lycopersicon esculentum cDNA clone

DEFINITION cTOF23N12 5' sequence, mRNA sequence.
ACCESSION BG129304 GI:12629492
VERSION 1
KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 630)

REFERENCE van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished (2001)

AUTHORS
TITLE Contact: CUGI
JOURNAL Clemson University Genomics Institute
COMMENT Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES
source
1..630
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOF23N12"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"

```
/lab_host="SOLR"
/clone_lib="tomato shoot/meristem"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

ORIGIN
Alignment Scores:
Pred. No.: 1.03e-92 Length: 630
Score: 1003.50 Matches: 196
Percent Similarity: 95.73% Conservat: 6
Best Local Similarity: 92.89% Mismatches: 6
Query Match: 28.16% Indels: 3
DB: 4 Gaps: 2

US-10-624-201A-2 (1-688) x BG129304 (1-630)

QY 451 TrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSer 470
DB 3 TGGAAAGCCCAATGGTTGAAGAAATGACTTGGGAAGAGTCAAGAAATCAAGAACAAACAGT 62
QY 471 ThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGlu 490
DB 63 AGTAAATACCTCAGGAGATACAAACAAAGAGACGATATAGTCTCCAAATGAAGAG 122
QY 491 LysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGlu 510
DB 123 AAACAACCAATTTATTACTAGCAGCTTTATTCAAGATGGT-----ACTACTCAAGCAGAA 176
QY 511 IleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHisAlaHis 530
DB 177 ATTTCTACTCTCAACTATTTCACCTTCCCTACTGCAGGTGCTTCACTTCATCGCTCAC 236
QY 531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIle 550
DB 237 AACTTCTCTCTCTGGTTCATTCAACATGAGAAATACTACTACTACTTGTGATCATATT 296
QY 551 GluAsnAsnAlaLysLysGlnArgAsn---AspMetHisLysPheSerProSerSerIle 569
DB 297 GAAACAAACGCGMAAAGCCAAAGAAATCATGACATGCACAAAGTTTCTCCAAAGTAGTATT 356
QY 570 LeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsn 589
DB 357 CTTTCATCTGTGAATGGAGCCAAAGCTAGAGAATCAACAAACAAAGGGTTTACTAAT 416
QY 590 ProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspProHisAspGlnGln 609
DB 417 CTTTAAATGGCAGCGTACCGATGGAGATTTTGGAGGTTTGTATCTTCATGATCAACAA 476
QY 610 MetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProProSer 629
DB 477 ATGACCGGGAATTTTCATGGAAATTAATGGTGTATCTCTTACTTTAGGACTTCTTCTCT 536
QY 630 GluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeuGlySerArg 649
DB 537 GAAACACCTAGCCATGCCAGTACCGATGGAGATTTTGGAGGTTTGTATCTTCTTAATGAGTATG 596
QY 650 SerGluMetGlySerHisTyrAsnArgMetGly 660
DB 597 CCTGAAATAGGAGTCAATACATAGAAATGGGA 629

RESULT 6
CO909070
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CO909070 610 bp mRNA linear EST 16-AUG-2004
BJ02039F01 BJ02 Capsicum annuum cDNA 5', mRNA sequence.
CO909070
CO909070.1 GI:51299373
EST.
KEYWORDS
SOURCE
ORGANISM

Capsicum annuum
Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

asterids; lamids; Solanales; Solanaceae; Capsicum.

1 (bases 1 to 610)

Song, W.Y., Jeon, W.B., Kim, K.S., Lee, H.H., Ko, M.K., Kim, Y.S., Hong, J.C. and Oh, B.J.

Generation of Hot Pepper (Capsicum annuum) ESTs (Express Sequence Tags) from Red Ripe Fruit (Song, et al.)

Unpublished (2004)

Contact: Boungh-Jun Oh

Jeonnam Biotechnology Research Center

Namyang Bld. #603, 10-4 Gwangsan-dong, Dong-gu, Gwangju 501-180, Korea

Tel: 82 62 607 2422

Fax: 82 62 607 6205

Email: bjo@biohub.re.kr

Plate: 039 row: F column: 01.

Location/Qualifiers

1. 610

/organism="Capsicum annuum"

/mol_type="mRNA"

/cultivar="Nokkwang"

/db_xref="taxon:4072"

/tissue type="red pepper fruit pericarp"

/clone_lib="BJ02"

/note="Vector: pBluescript II SK(+); Site_1: EcoRI; Site_2: XhoI; cDNA library was generated from red ripe fruit pericarp using lambda Zap II phage vector. In vivo excision was done with helper phage to generate subclone in pBluescript II SK(+); vector."

ORIGIN

Alignment Scores:

Pred. No.: 1.82e-90 Length: 610

Score: 981.50 Matches: 192

Percent Similarity: 95.63% Conservat: 5

Best Local Similarity: 93.20% Mismatches: 6

Query Match: 27.55% Indels: 3

DB: 7 Gaps: 1

US-10-624-201A-2 (1-688) x CO909070 (1-610)

QY 250 ValAsnThrAsnSerSerGlyGlyGluSerSerArgGlnLysAsnGluValAla 269

DB 1 GTCACACTAGTAGT-----GGTACTGGTAGCAGCGCCGAAATAAGATTGCT 51

QY 270 ValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAla 289

DB 52 GCTGAGCTTCAACTGCTCAAGACAGAACTTCAGATGAAGAAAGGCAAGCTTGTACC 111

QY 290 MetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleVal 309

DB 112 ATGCTTGAAGAGGTGGAACAAAGGTACAGACAGTACCATCAAAATGCAATTAATTGTG 171

QY 310 LeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeu 329

DB 172 TCATCATTTGAGCAAGTAGCAGGAATTTGATCAGCAAAATCATACATCAACTGCTTGTG 231

QY 330 HisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAla 349

DB 232 CATCAATTTGGAAGCAATTCGATGCTAAAGATGCAATTCATGCAACAAATTAAGGG 291

QY 350 ThrSerLysSerLeuGlyGluGluGluGlyLysIleGluGlySerArgLeu 369

DB 292 ACAGAGCAAGAGTTTAGGTGAAGAGGAGGATAGGAGGAGGAGGAGGAGGAGGAGGAGG 351

QY 370 LysPheValAspHisHisLeuArgGlnGlnAlaLeuGlnGlnIleGlyMetMetGln 389

DB 352 AAATTTGTAGACCATCATCTAAGGCAACACAGTGGCTGCAACAAATTAGGAATGATGCA 411

QY 390 ProAsnAlaTyrArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAla 409

DB 412 CCAATGCTTGGAGACCTCAAGAGAGGTCTACTCAAGAGGGGTGTCCTCGCTCTCGTCT 471

QY 410 TrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAla 429

```

472 TGGCTTTTCGAGCATTTTCTTCATCCTTACCCCAAGGATTCAGACAAAATCATGCTTGCT 531
Db
430 LysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArg 449
Qy
532 AAGCAAAACCGGACTAAACAAGGAGCGAGGTCTCCAACTGGTTTCATAAAATGCTCGAGTTTCCA 591
Db
450 LeuTrpLysProMetVal 455
Qy
592 CTATGGAAGCAATGGTG 609
Db

RESULT 7
BQ119770 699 bp mRNA linear EST 07-MAR-2003
LOCUS EST605346 mixed potato tissues solanum tuberosum cDNA clone STNEM54
DEFINITION 5' end, mRNA sequence.
BQ119770
ACCESSION BQ119770.2 GI:21919664
VERSION BQ119770
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 699)
AUTHORS Buell,C.R., Hart,A., Baker,B., Tankalev,S., Fry,W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karamycheva,S.A.
TITLE Generation of a set of potato cDNA clones for microarray analyses
JOURNAL Unpublished (2002)
COMMENT On Apr 17, 2002 this sequence version replaced gi:20171732.
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.
FEATURES
Location/Qualifiers
source
1..699
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Binjete"
/db_xref="taxon:4113"
/clone="STNEM54"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="mixed potato tissues"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating
eyes, tubers, or roots."
ORIGIN
Alignment Scores:
Pred. No.: 8,99e-89 Length: 699
Score: 966.00 Matches: 183
Percent Similarity: 99.46% Conservative: 0
Best Local Similarity: 99.46% Mismatches: 1
Query Match: 27.11% Indels: 0
DB: 5 Gaps: 0

US-10-624-201A-2 (1-688) x BQ119770 (1-699)
Qy 1 MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHis 20
Db 148 ATGTACTATCAAGGACCTCGGTAATACTAATATACAAAGCTGATCAACAACATCAT 207
Qy 21 AsnHisGlyAsnSerAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
Db 208 AATCATGGGAATAGTAATAATAATAATATTTCAGACACTTTTATTTTATGATGAACCTTAACAAT 267

```

```

Qy 41 TyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnLeuLeuPheLeuAsn 60
Db 268 TATATGCAAGGCTACACTTCTTGACACACAGCAGCAGCAGTGTACTTTTCTCTGAAT 327
Qy 61 SerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGln 80
Db 328 TCTTCCACCAGCAGCAAGCAAGCGCTTTGCCATGCGAATATACAACACGCGCGCTCAA 387
Qy 81 GlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsnHis 100
Db 388 CAGCAGCACTTTGTGCGGTGTGCCTCTTCCGCGCAGTAGTTTGCACGATCATCAATCAT 447
Qy 101 HisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValPro 120
Db 448 CATGGACTTTTACAGCGCATGTGGACACCAAGATCAATCTCAGCAGGTGATAGTACCA 507
Qy 121 SerSerThrGlyValSerAlaThrSerCysGlyIleThrThrAspLeuAlaSerGln 140
Db 508 TCGTCGACGGGGTTTCTGCCACGTCATGTGGCGGATCACACGGGACTTTGGCGTCTCAA 567
Qy 141 LeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGly 160
Db 568 TTGCGCGTTTCAGAGCGCGATTTCGACACCAACACACGACGACGACCAACAGCAAGGC 627
Qy 161 GlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIle 180
Db 628 GGTCTATCTCTAAGCCCTTTCTCTCAGCTACACAGCAAAATTAGTTTCAATAACAATATT 687
Qy 181 SerSerSerSer 184
Db 688 TCATCCTCATCA 699

RESULT 8
CL978605 1965 bp DNA linear GSS 21-SEP-2004
LOCUS OaIFCC032180 Oryza sativa Express Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL978605
VERSION CL978605.1 GI:52411712
KEYWORDS GSS.
ORGANISM Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 1965)
AUTHORS Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
Location/Qualifiers
source
1..1965
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/notes="Oryza sativa exon trapped genomic sequences "
ORIGIN
Alignment Scores:
Pred. No.: 8,35e-88 Length: 1965

```

[illegible]

745	CGGCAGATGCAGGGGTGGCGCGCGTTCGAGCGCGCGCGGGCGCGGTTCGGCGACG	804
323	SerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAla	342
805	ACGTACACGTCTGGCGCTGCCGACCATCTCCGCGCAGTTCCGGTGCCTCCCGCAGCC	864
343	IleAlaGlnValLysAlaThrSerLysSerLeuGlyGlu	358
865	ATCGCGCGCAGGTACGGCGCGCAGCCGGGGCTCGGGGAGGACTGCGCGCAGCAGGAG	924
359	GlyLeuGlyGly	375
925	GGCGCGGTGGCGCGGAGGACGACGCTGGGTTCGCGCTTCGCGGTTCATCGACACCAG	984
376	LeuArgGlnArgAlaLeuGlnGlnIleGlyMetMetClnProAsnAla	392
985	CTCCGCGCAGCGCGCGATGTCAGCAGCTTGGCATGTGTGCACGCCCGCGCGGTGGC	1044
393	-----TtpArgProGlnArgGlyLeuProGlnArgAlaValSerValLeu	407
1045	CGCGCGCGCGGGTGGCGCGCGCAGCGTGGCTCTCCCGAGCGCGCGCTCTCCGTCTC	1104
408	ArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMet	427
1105	CGCGCTGGCTCTTCAGCATTTCTCCACCCATACCCCAAGGATTCGACAAAGTTCATG	1164
428	LeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArg	447
1165	CTGCCAAGCAACCGCGCTCACCGAGCCAGGTGTGAACTGTTCATCAACGCGAGG	1224
448	ValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGlu	467
1225	GTGAGGCTGGGAAGCCATGGTGGAGGAGATGTACCGAGGAGACCAAGGCCAAGGAG	1284
468	GlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaPro	487
1285	GAG-----	1287
488	AsnGluGluLysHisProIleThrSerSer-----LeuLeuGln	501
1288	GAGGAGGAGGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAG	1347
502	AspGlyIleThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerProThr	521
1348	CAGGCTCCGAGTAAGCCGCGCAGCAGTGTGGCATTTGGCATTCGTCGTCGTCGCGCGG	1407
522	AlaGlyAlaSer-----LeuHis-----HisAlaHisAsnPheSer	533
1408	CGCGCGCGCTCCAGGTCAAGTTGGTGTGTGCACCGCGGAGACGAGCAGCTCAGGCGG	1467
534	PheLeuGlySerPheAsnMetAspAsn-----ThrThr	544
1468	TACCGG	1527
545	ThrThrValAsp-----HisIleGluAsnAsnAlaLys---	555
1528	ACAGCGCGGAGCAGCGG	1587
556	LysGlnArgAsnAspMetHisLysPheSerProSerSer-----	568
1588	AGCCACCGGAGCTGCTAATGAAGTTCACTGAGCGCGCGCGCGCGCGCGCGCGCGG	1647
569	-----IleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGly	586
1648	CACCCGACGCTCAACGACGACGACGACGCTCCCTGGCGCGCGCGCGCGCGCGCGCTG	1707
587	PheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspProHis	606
1708	TTACG-----GCGGCGCAGTACGGGCATCAGTTTCGGC-----TCGGACCAT	1749
607	AspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeu	626

Db 1750 TTCGCGTTCGCGGCACGCGCGGAGCGCGCGCGTGTGCTCAGCTCGCGCTC 1809
Qy 627 ProProSerGluAenLeuAlaMetProValSerGlnGlnAenTyrLeuSerAsnAspLeu 646
Db 1810 CCGCAGCGCGCGAC---CAGACGCGCGCTCG-----TTCATAGGCGCGGT 1857
Qy 647 GlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGluAenIleAspPhe 666
Db 1858 GCCGGCAGCAGCGCGCGCGCGCGGTGTCAGCAGCGCGGTACGAC---ATGAACATG 1914
Qy 667 GlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
Db 1915 CAGAGC---ACCAAGTCGTTGGCTGCTCAGCTCATGAGAGACTTCGTG 1959

RESULT 9
BI920940
LOCUS BI920940 708 bp mRNA linear EST 10-MAR-2003
DEFINITION EST540875 potato microtubers, in vitro-grown Solanum tuberosum cDNA
clone cSTE26N18 5' end, mRNA sequence.

ACCESSION BI920940
VERSION BI920940.1 GI:16216968
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 708)
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R.,
Karanycheva,S.A., Tsai,J., Van Aken,S., Utterback,T., Chienmingo,A.,
Bougrl,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generation of ESTs from in vitro grown microtubers (2001b)
Unpublished (2001)
CONTACT: Robin Buell

COMMENT The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.

FEATURES
source
1. 708
Location/Qualifiers
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="cSTE26N18"
/tissue_type="axillary buds of stem explants; growing
sink-tubers"
/dev_stage="7, 8 and 10 days"
/lab_host="SOLR"
/clone_lib="potato microtubers, in vitro-grown"
/note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Tissue supplied by Christian Bachem and Richard
Visser (Department of Plant Breeding, Wageningen
University, The Netherlands); sequencing by The Institute
for Genomic Research. The cSTA libraries will attempt to
capture the induction and initiation/initial growth of the
tuber in an in vitro system as described in Bachem et al.
(Plant Journal, 1996). Small microtubers develop from
axillary buds attached to stem explants when placed on a
high sucrose medium (10%). Visible morphological changes
occur synchronously at day five in the axillary buds. The
first library, cSTA (1-20) consists of axillary buds
harvested on days 1-3. This targets those genes involved
in induction of the microtubers. The following libraries,
cSTA (21-40) and cSTA (41-60), capture genes involved in
tuber initiation and outgrowth. This library is noted as
P3 in Tanksley lab notebooks."

ORIGIN

Alignment Scores:

Prod. No.: 3.01e-88 Length: 708
Score: 961.00 Matches: 186
Percent Similarity: 94.97% Conservative: 3
Best Local Similarity: 93.47% Mismatches: 8
Query Match: 26.97% Indels: 2
DB: Gaps: 2
US-10-624-201A-2 (1-688) x BI920940 (1-708)
Qy 1 MetTyrTyGlnGlyThrSerAspAenThrAsnIleGlnAlaAspHisGlnGlnArgHis 20
Db 110 ATGTACTATCAGGAACCTCGGATAAT---AATATCAAGCTGATCATCAACACATCAT 166
Qy 21 AsnHisGlyAsnSerAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
Db 167 AATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTGTGATGAACCTTAACAAT 226
Qy 41 TyrMetGlnGlyTyrThrThrSerAspThr---GlnGlnGlnGlnGlnLeuLeuPheLeu 59
Db 227 TATATGCAAGGCTACACTACTCTTGACACACAGCAGCAGCAGCAGGTTACTTTTCTGT 286
Qy 60 AsnSerSerProAlaAlaSerAsnAlaLeuCybHisAlaAsnIleGlnHisAlaProLeu 79
Db 287 AATCTTCACACAGAGAAACGCGCTTTGCCATCGGAATATACAAACGCGCCCTG 346
Qy 80 GlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsn 99
Db 347 CAACAGCAGCATTCTTGGTGTGCTCTTCGCGCAGTAGTAACTTGCACGATCAGATCAAT 406
Qy 100 HisHisGlyLeuLeuGlnArgMetTyrAsnAsnGlnAspGlnSerGlnGlnValIleVal 119
Db 407 CATCATGGACTTTTACAGCGTATGTGGAACAACCAAGATCAATCTCAGCAGGTTAGTAGTA 466
Qy 120 ProSerSerThrGlyValSerAlaThrSerCybGlyGlyIleThrThrAspLeuAlaSer 139
Db 467 CCATCGTCGACGGGGTTTCTGCCAGCTCATGTGGCGGATTAACCGAGCTTGGCGTCT 526
Qy 140 GlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGln 159
Db 527 CAATTGGCGTTTCAGAGCGCGATTGTGGTTCGGAACACCCGACGACGACCAACACAA 586
Qy 160 GlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsn 179
Db 587 GCGCGTCTATCTCTAAGCCTTTCTCCTCAGCAACAACAGCAAAATTAAGTTTCAATAACAAT 646
Qy 180 IleSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGly 198
Db 647 ATTTGAGCTTCATCACCACCAAGGACAATAATGTACTATTAGAGAAAGATTAGATGGG 703

RESULT 10
AW441343
LOCUS AW441343 623 bp mRNA linear EST 18-MAY-2001
DEFINITION EST310739 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
clone cLEN15B11 5', mRNA sequence.

ACCESSION AW441343
VERSION AW441343.1 GI:6976594
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum

REFERENCE 1 (bases 1 to 623)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,P.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
CONTACT: CUGI

COMMENT Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.
Location/Qualifiers

FEATURES
source

1..623
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEN15B11"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_lib="tomato fruit red ripe, TAMU"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

ORIGIN

Alignment Scores:
Pred. No.: 3,896-85 Length: 623
Score: 930.00 Matches: 177
Percent Similarity: 94.55% Conservative: 14
Best Local Similarity: 87.62% Mismatches: 9
Query Match: 26.10% Indels: 2
DB: 2 Gaps: 1

US-10-624-201A-2 (1-688) x AW441343 (1-623)

Qy 270 ValGluLeuThrAlaGlnArgGlnGluLeuMetLysAlaLysLeuAla 289
Db 8 GTTGAACTTACACAGCTCAAGACAGAACTTCAATGAGAGCAAGCTTGTAGC 67
Qy 290 MetLeuGluGluValGluGlnArgTyrArgGlnTyrHisGlnMetGlnIleVal 309
Db 68 ATGCTTGATGAGTCAAGGTTATAGACAAATACCAATCAATGCAATGATTGCA 127
Qy 310 LeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAla 329
Db 128 ACATCATTTGAGCAACACAGGAATTTGGATCATCAATCAATCATACACAACTTGTCTTG 187
Qy 330 HisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAla 349
Db 188 CACACAAATTTCAAGCAATTTAGATGTTTAAAGATGCAATTTCTGGGCAATTAAGGAC 247
Qy 350 ThrSerLysSerLeuGlyGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeu 369
Db 248 ACAATAAACTTTAGGGGAAGAGAGAACATTTGGTGGCAAAATTTGAAGGATCAAGTTG 307
Qy 370 LysPheValAspHisGlnArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln 389
Db 308 AAATTTGGGATCATCATTTAGCCCAACACAGTGCATACCAATTAAGGATGATGCA 367
Qy 390 ProAenAlaTpaArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAla 409
Db 368 ACCAATGATGAGGCCCAAGAGGTTTGGCCGAAGAGGGTTTCGGTTCTCGCGCT 427
Qy 410 TrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAla 429
Db 428 TGGCTTTTCGACATTTCTTCATCCGTATCTCTAAAGATTCAGATAAATCATGCTTCT 487
Qy 430 LysGlnThrGlyLeuThrArgSerGlnValSerAenTrpPheIleAsnAlaArgValArg 449
Db 488 AAGCAACAGGGCTAACAGGAGGACAGGTATCAAAATTTGGTTTCATAAATGCTAGAGTTAGA 547
Qy 450 LeuTrpLysProMetValGluGluMetTyrLeuGluGluValLys-----AsnGlnGlu 467
Db 548 CTATGGAAGCCCAATGGTAGAAGAAATGTACATGGAAGAAGTGAAGAAATACAAATCAAGA 607

Qy 468 GlnAen 469

Db 608 CAAAAT 613

RESULT 11

AI773253

LOCUS

DEFINITION

EST254353 tomato resistant, Cornell Lycopersicon esculentum cDNA

clone CLERSL8, mRNA sequence.

AI773253

ACCESSION

VERSION

AI773253.1 GI:5271294

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum (tomato)

Eukaryota; Viridiplantae;

Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Lycopersicon.

1 (bases 1 to 553)

REFERENCE

AUTHORS

D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,

Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,

Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,

Giovannoni, J. and Martin, G.B.

Generation of ESTs from Pseudomonas resistant tomato

Unpublished (1999)

CONTACT: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

Location/Qualifiers

1..553

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"

/db_xref="taxon:4081"

/clone="CLERSL8"

/tissue_type="leaf"

/dev_stage="4-week old"

/lab_host="SOLR"

/clone_lib="tomato resistant, Cornell"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; CLER - Tomato Pseudomonas Resistant EST Library.

Directionally cloned cDNAs inserted into pBluescript

SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN

Alignment Scores:

Pred. No.: 3,476-84 Length: 553

Score: 920.00 Matches: 175

Percent Similarity: 98.36% Conservative: 5

Best Local Similarity: 95.63% Mismatches: 3

Query Match: 25.82% Indels: 0

DB: 1 Gaps: 0

US-10-624-201A-2 (1-688) x AI773253 (1-553)

Qy 317 GlyTleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPhe 336

Db 3 GGAGTTGATCAGCAAAATCATACACTTATGCTTTCATGCAATTTCAAGCAATTC 62

Qy 337 ArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu 356

Db 63 AGATGCCTAAAGGATGCAATTTAGTGAGCAAGTGAAGCGACGACGAGTGTAGTGCA 122

Qy 357 GluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisLeu 376

Db 123 GATGAAGCTTTGGGAGGGGAAAATTTGAAGGCTCAAGACTCAAAATTTGTGGACCATCATCTA 182

Qy 377 ArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAenAlaTpaArgProGln 396

Db 183 AGGCAACACCGCGCTGCAACAGCTAGGAATGATGCAACCAATGCTTGGAGACCCCA 242

```
Qy 397 ArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeu 416
Db 243 AGAGTTTACTCCTGAAGAGCTGTCTGTCTTGGCTTTTCGAGCATTTTCTT 302
Qy 417 HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArg 436
Db 303 CATCTCTTACCCCAAGACTCAGCAAAATCATGCTTCTAAGCAACCGGGCTTAACAAG 362
Qy 437 SerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGlu 456
Db 363 AGCAGGTCTCTAACTGGTTCATTAACTCGAGTTCGATTATGGAAGCCAAATGGTTGAA 422
Qy 457 GluMetTyrLeuGluGluValLysAsnGlnGlnAsnSerThrAsnThrSerGlyAsp 476
Db 423 GAAATGTACTTGGGAAGTGAAGAATCAAGAACAAACAGTAGTAGTAATCTTCAGGAGAT 482
Qy 477 AsnLysAsnLysGluThrAsnIleSerAlaProAsnGluLysHisProIleIleThr 496
Db 483 AACAAAAACAAGAGACGAAATTTAGTGTCTCCAAATGAAGAGAAACAACCCATTATTACT 542
Qy 497 SerSerLeu 499
Db 543 AGCAGCTTA 551

RESULT 12
CV241175/c
LOCUS
DEFINITION CV241175.1 827 bp mRNA linear EST 22-SEP-2004
CDNA clone WS02511_G23 3', mRNA sequence.
ACCESSION CV241175
VERSION CV241175.1 GI:52494150
KEYWORDS EST.
SOURCE
ORGANISM
Populus balsamifera subsp. trichocarpa (Populus trichocarpa)
Populus balsamifera subsp. trichocarpa
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
1 (bases 1 to 827)
REFERENCE
Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
Bohlmann,J.
The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)
Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-6097
Email: bohlmann@interchange.ubc.ca
Plate: WS02511 row: G column: 23
High quality sequence stop: 827.
FEATURES
Location/Qualifiers
1..827
/organism="Populus balsamifera subsp. trichocarpa"
/mol_type="mRNA"
/cultivar="Wild clone"
/sub_species="trichocarpa"
/db_xref="taxon:3694"
/clone="WS02511_G23"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PT-MB-N-A-15"
/notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5'
end of cDNA); Site 2: XhoI (3' end of cDNA); Terminal
vegetative buds from 20 year old trees harvested near
Corvallis, Oregon on September 19th, 2001. cDNA was
prepared from 5 micrograms of mRNA and directionally
```

ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldi M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

ORIGIN

```
Alignment Scores:
Pred. No.: 1,22e-81 Length: 827
Score: 898.00 Matches: 184
Percent Similarity: 73.72% Conservative: 32
Best Local Similarity: 62.80% Mismatches: 55
Query Match: 25.20% Indels: 22
DB: 7 Gaps: 4

US-10-624-201A-2 (1-688) x CV241175 (1-827)
Qy 155 GlnGlnGlnGlnGlnGlyLeuSerLeuSerProGlnLeuGlnGlnIle 174
Db 820 CAGGCTCAAACTGTATCTGTGTAAGATATTAAGGGTGTC----- 782
Qy 175 SerPheAsnAsnAsnIleSerSerSerProArgThrAsnAsnValThrIleArgGly 194
Db 781 -----GGTGGGTCGGGTGCATCAGGCTCGGGTGTGACAAATGGGGTCTCTG----- 737
Qy 195 ThrLeuAspGlySerSerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGln 214
Db 736 -----GGTATGCAAGGGGTGTGCTGAGCTCAAAAGTACTTTGAAGGGTCTCAA 689
Qy 215 GluLeuLeuAspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLys 234
Db 688 GAGTACTTGTATGAGGTTGTTCGGCGAATAATAATGATATTAGAGTGAATTCGAAG 629
Qy 235 LysAspAsnSerMet-----AsnLysGluSerMetProLeuAlaSerAspValAsnThr 252
Db 628 AGGAGTAATGGGATTGGTAGTAACACTAGTAACAAAGTGGTTGGAGAATCATTAGCCGGA 569
Qy 253 AsnSerSerGlyGlyGlyGluSerSerArgGlnLysAsnGluValAlaValGluLeu 272
Db 568 GAAAGGGTCTCGTGGAGAGAAGTAAGTGGGAAGCGT-----GGACCGGAGCTT 521
Qy 273 ThrAlaGlnArgGlnGluLeuMetLysLysAlaLysLeuLeuAlaMetLeuGlu 292
Db 520 TCCACTGCCGAGAGCGCAGGAAATTCAGATGAAGAAGCTAAACTTATCAGCATGCTTGAT 461
Qy 293 GluValGluGlnArgTyrArgGlnTyrHisGlnMetGlnIleIleValLeuSerPhe 312
Db 460 GAGGTGGAGCAAGGTACAGGCAGTATCATCACCAGATGCAGATTGTGATTTCCTCGTT 401
Qy 313 GluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIle 332
Db 400 GAGCAAGCAGTAGGAATTTGGTTCGCTTAAGCATATATACAGCCCTTGGCGTGAACAATC 341
Qy 333 SerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLys 352
Db 340 TCCAAGCAGTTTGGTGCTTGAAGGATGCAATTAACAGTCAAATTAAGCTGCAACAA 281
Qy 353 SerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheVal 372
Db 280 AGCTTAGTGAAGAGGATTGCTTGGGAGGAAGATTGAAGGTTCAAGGCTCAAAATTGTC 221
Qy 373 AspHisIleLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAla 392
Db 220 GATCATCACCTTAGGCAACAGCGTGCATTTCAAGGTTTGGGAATGATCCAGCACAATGCT 161
Qy 393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db 160 TGGAGACACAGAGAGATTGCTTGAAGAAATTCAGTTTCACTTCTCCGTGCTTGGCTCTTC 101
```

```

Qy 413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db 100 GAACACTTTCACCCCTATCCCAAGGATTGACACAAGCAGCATGCTCGCAAAACAACA 41
Qy 433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsn 445
Db 40 GGACTCAGGAGGCCAGGTGCTAAATTTGGTTCTATAAAC 2

RESULT 13
LOCUS B1921008 789 bp mRNA linear EST 10-MAR-2003
DEFINITION EST540943 potato microtubers, in vitro-grown Solanum tuberosum cDNA
clone cSTE27121 5' end, mRNA sequence.
ACCESSION B1921008
VERSION B1921008.1 GI:16217036
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 789)
van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R.,
Karameycheva,S.A., Tsai,J., Van Aken,S., Utterback,T., Chiemingo,A.,
Bongri,O., Buell,C.R., Ronning,C., Tankalev,S. and Baker,B.
Generation of ESTs from in vitro grown microtubers (2001b)
Unpublished (2001)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.

FEATURES
    source
        Location/Qualifiers
            1..789
                /organism="Solanum tuberosum"
                /mol_type="mRNA"
                /cultivar="Bintje"
                /db_xref="taxon:4113"
                /clone="cSTE27121"
                /tissue_type="axillary buds of stem explants; growing
                sink-tubers"
                /dev_stages="7, 8 and 10 days"
                /lab_host="SOLR"
                /clone_lib="potato microtubers, in vitro-grown"
                /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                XhoI; tissue supplied by Christian Bachem and Richard
                Visser (Department of Plant Breeding, Wageningen
                University, The Netherlands); sequencing by The Institute
                for Genomic Research. The cDNA libraries will attempt to
                capture the induction and initiation/initial growth of the
                tuber in an in vitro system as described in Bachem et al.
                (Plant Journal, 1996). Small microtubers develop from
                axillary buds attached to stem explants when placed on a
                high sucrose medium (10%). Visible morphological changes
                occur synchronously at day five in the axillary buds. The
                first library, cSTA (1-20) consists of axillary buds
                harvested on days 1-3. This targets those genes involved
                in induction of the microtubers. The following libraries,
                cSTA (21-40) and cSTA (41-60), capture genes involved in
                tuber initiation and outgrowth. This library is noted as
                P3 in Tanksley lab notebooks."

ORIGIN
Alignment Scores:
Pred. No.: 2,31e-81 Length: 789
Score: 895.00 Matches: 168
Percent Similarity: 99.41% Conservative: 0
Best Local Similarity: 99.41% Mismatches: 1
Query Match: 25.12% Indels: 0
Db: 4 Gaps: 0

```

```

US-10-624-201a-2 (1-688) x B1921008 (1-789)
Qy 1 MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnArgHis 20
Db 282 ATGTACTATCAAGGAACCTCGGATAACTATAATACAAAGCTGATCATCAACAATCAT 341
Qy 21 AsnHisGlyAsnSerAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
Db 342 AATCATGGGAATAGTAATAATAATATTTCAGACACTTTATTTTGTGAAACCTTAACAAT 401
Qy 41 TyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGlnLeuLeuPheLeuAsn 60
Db 402 TATATGCAAGGCTACACTTCTTGACACACAGCAGCAGCAGCAGTACTTTTCTCTGAAT 461
Qy 61 SerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGln 80
Db 462 TCTTCACCAGCAGCAAGCAGCGCTTTGCCATCGGAATATACAACACGCGCGCTGCAA 521
Qy 81 GlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsnHis 100
Db 522 CAGCAGCACTTTTGTGGTGTGCCTCTTCGCGCAGTAAGTTTGCACGATCAGATCAATCAT 581
Qy 101 HisGlyLeuLeuGlnArgMetTyrAsnAsnGlnAspGlnSerGlnGlnValIleValPro 120
Db 582 CATGGACTTTTACAGCGCATGTGGAAACACCAAGATCAATCTCAGAGGTGATAGTACCA 641
Qy 121 SerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGln 140
Db 642 TCGTCGAGCGGGGTTTCTGCCACGTCATGTGGCGGGATCACCCAGGACTTCGCGTCTCAA 701
Qy 141 LeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGly 160
Db 702 TTGGCGTTTTCAGAGCGCGATTCCGACACCAACACCGACAGCAGCAACAACAGAGGCG 761
Qy 161 GlyLeuSerLeuSerLeuSerProGln 169
Db 762 GGTCTATCTCTAAGCGCTTTTCTCCTCAG 789

RESULT 14
LOCUS CB292855 850 bp mRNA linear EST 28-FEB-2003
DEFINITION UCRCS01_05af07 b1 Washington Navel orange cold acclimated flavedo &
albedo cDNA library Citrus sinensis cDNA clone UCRCS01_05af07, mRNA
sequence.
ACCESSION CB292855
VERSION CB292855.1 GI:28618312
KEYWORDS EST.
SOURCE Citrus sinensis
ORGANISM Citrus sinensis
REFERENCE 1 (bases 1 to 850)
AUTHORS Close,F.J., Collura,K., Fenton,R.D., Feuerbacher,O., Kim,H.R.,
Kudrna,D., Wananaker,S., Wing,R. and Yu,Y.
DEVELOPMENT Development of EST Resources and New Genetic Markers for California
Citrus
CONTACT Contact: Timothy Close
DEPARTMENT Department of Botany & Plant Sciences, University of California
RIVERSIDE, CA, 92521-0124
TEL: 9097873318
FAX: 9097874437
EMAIL: timothy.close@ucr.edu
Seq primer: T7.
FEATURES
    source
        Location/Qualifiers
            1..850
                /organism="Citrus sinensis"
                /mol_type="mRNA"
                /cultivar="Washington navel"
                /db_xref="taxon:2711"
                /clone="UCRCS01_05af07"

```

/tissue_type="Rind containing flavado and albedo"
/dev_host="Mature fruit"
/lab_host="E. coli TJCU21"
/clone_lib="Washington Navel orange cold acclimated
flavado & albedo cDNA library"
/notes="vector: lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
field at University of California, Riverside Agricultural
Operations since 1983. The scion was Washington Navel
orange and the rootstock Carizzo Citrange. Tissue from
mature fruit was harvested at mid-day in January 2002
during a cold spell, when pre-dawn temperatures were
approximately -2 to 2 degree C. Approximately 2 cm median
sections of the rind were excised in the field from
several fruits, then wrapped in aluminum foil and frozen
quickly in dry ice. Total RNA was extracted using a phenol
extraction procedure described in J. Japanese Soc. Hort.
Sci. (1996) 64:809-814. Poly(A) RNA was purified, a cDNA
library was made, and 1 million primary lambda cDNA clones
were in vivo excised to give a population of pBluescript
SK(-) phagemids. All steps to this point were performed in
the TJ Close lab at the University of California,
Riverside (Fenton). Phagemids were plated, plasmid DNA
purified, cDNA clones archived, and DNA sequences
determined bi-directionally using an ABI3530 at the
Arizona Genomics Institute, University of Arizona
(Collura, Feuerbacher, Kim, Kudrna, Wing, Yu).
Chromatogram files were transmitted to UC Riverside (by
Yu), then processed at UC Riverside (by Wanmaker) using
the Harvest pipeline (<http://harvest.ucr.edu>) to remove
vector and cloning oligo sequences and various
contaminants, and to trim to a high quality region.
Sequences that retained a phred 17 region of at least 100
bases were deposited to GenBank."

ORIGIN

Alignment Scores:
Pred. No.: 4,71e-81 Length: 850
Score: 892.50 Matches: 183
Percent Similarity: 73.59% Conservative: 26
Best Local Similarity: 64.44% Mismatches: 50
Query Match: 25.05% Indels: 25
DB: 6 Gaps: 3

US-10-624-201A-2. (1-688) x CB292855 (1-850)

Qy 272 LeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeu 291
Db 17 CTCTCCAGCGCAGAGAGCGCAAGAAATTCAGATGAAGAAAGCTCATTAAAGTAATATGCTT 76

Qy 292 GluGluValGlnArgGlnTyrArgGlnTyrHisGlnMetGlnMetGlnLeuValLeuSer 311
Db 77 GATGAGTGGAGCAGAAAGTACCCGTTGAACCATCACCAGATGCAGATGATGATTCATCA 136

Qy 312 PheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAla 331
Db 137 TTGTGAACAGCAGCAGCAAGATTTAGTCAAGCAAGACATACACAGCCTTCTCTCAAGACA 196

Qy 332 IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer 351
Db 197 ATTTCAAGCAATTCGCGTCTGAAAGAGTCAATTCAGGCCCAAAATTAAGCGGCAAAAT 256

Qy 352 LysSerLeuGlyGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPhe 371
Db 257 AAGATGCTGGGTGAGGAGGATTCGCTCGGAGCAGAGATTGAAGCTCCAGGCTCAAAATTT 316

Qy 372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
Db 317 GTCGACCACCATCTTCGTCAACAAAGGCTCTTCAACAAATTTGGGAATGATCCAGCACAAT 376

Qy 392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
Db 377 GCTTGGAGAGCCCCAAGAGGATTCCTGAAAGATCTCTCTCAGTCTTCGCGCTTGGCTG 436

Qy 412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
Db 437 TTCGAACACTTCCTTCCACCGTATCCGAAAGATTCCGACAAGCAGATGCTGGCAAAACA 496

Qy 432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
Db 497 ACAGGGCTTACCAGAGCCAGGTTGTCTAAATTGGTTCTAATATCCAGAGTTCGCGTGTGG 556

Qy 452 LysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsn----- 469
Db 557 AAACCAATGTTGAGGAATGTACTTGGAGGAAATTAAGATCAAGAACAGACGAGTCA 616

Qy 470 -----SerThrAsnThrSerGly 475
Db 617 GAGGACAAACAGCAAGAGTGAGCAATGAGGATTTCGGCATCTAAGTCCATCGCTGCA 676

Qy 476 AspAsnLysAsn-----LysGluThrAsnIleSerAlaProAsnGluGluLysHisPro 493
Db 677 CAAGAGAAAAATCTCGTGAAGGAACTCAAAACTCGAAGAGTTTCAAAATCCAGTGGAGAT 736

Qy 494 IleIleThrSerSerLeuLeuGlnAspGlyIleThrThrGlnAlaGluIleSerThr 513
Db 737 AATTGACAAACAG-----AATGTTCAATCTATGATTTCAATG 775

Qy 514 SerThrIleSerThrSerProThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSer 533
Db 776 TCACAATTCACAACATCTCTATCGGAGGAGGAATCGGAGAACCATTCAGGATTTTCA 835

Qy 534 PheLeuGlySer 537
Db 836 CTCATTGGATCA 847

RESULT 15
BMS35737

LOCUS BMS35737 752 bp mRNA linear EST 20-FEB-2002
DEFINITION BMS35737 tomato breaker fruit Lycopersicon esculentum cDNA clone
cLEG67P6 5' end, mRNA sequence.
ACCESSION BMS35737
VERSION BMS35737.1 GI:18814614
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 752)
Alcala, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S.A.,
Tsai, J., Bougri, O., Kirkness, E., Utterback, T., Van Aken, S.,
Ronning, C.M., Fraser, C.M., Martin, G.B., Tankale, S.D. and
Giovannoni, J.
Generation of ESTs from tomato fruit tissue, breaker stage (2002)
Unpublished (2002)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
1. 752
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG67P6"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_lib="tomato breaker fruit"
/notes="Vector: pBluescriptSKCudapt; Site_1: EcoRI;

Site 2: XhoI: supplier: Boyce Thompson Institute;
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

ORIGIN

Alignment Scores:

Pred. NO.:	7.91e-81	Length:	752
Score:	889.50	Matches:	188
Percent Similarity:	82.47%	Conservative:	19
Best Local Similarity:	74.90%	Mismatches:	27
Query Match:	24.96%	Indels:	18
DB:	4	Gaps:	4

US-10-624-201A-2 (1-688) x BM535737 (1-752)

QY	237	AsnSerMetAsnLysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGly	256
DB	7	CACAAATTGGAGAAATTAATATCTTTGGCTAGTGAT-----	45
QY	257	GlyGlyGluSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGln	276
DB	46	---GATATTGAAAGTAACACCCCAAAAATAGT---GGTGTGAACTTACACAGCTCAA	99
QY	277	ArgGlnGluLeuGlnMetLysLysAlaLysLeuAlaMetLeuGluGluValGluGln	296
DB	100	AGACAAGAACTTCAAAATGAAGAAACAAAGCTTTAGCATGCTTGATGAGGTGGATCAA	159
QY	297	ArgTyrArgGlnTyrHisGlnMetGlnIleValLeuSerPheGluGlnValAla	316
DB	160	AGGTATAGCAATACCATCAACAAATGCAATGATTCGAACATCATTTGAGCAACACA	219
QY	317	GlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPhe	336
DB	220	GGATTTGGATCATCAAAATCATACACAACTTGCTTTGCACACAAATTTCAAGCAATTT	279
QY	337	ArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeu-GlyGln	356
DB	280	AGATGTTTAAAGATGCAATTTCTGGGCAATTAAGGACACAAATAAACTTTAGGGGGA	339
QY	356	uGluGluGlyLeuGlyLysIleGluGlySerArgLeu-LysPheValAspHisL	376
DB	340	AGAAGAGAACATTTGGTGGGCAATTTGAAGGATCAAAAGTTGAAATTTGTGGATCATCAT	399
QY	376	euArgGlnGlnArgAlaLeuGlnIleGlyMetMetGlnProAsnAlaTrpArgProG	396
DB	400	TACGCCAACACACGTGCACTACACAAATTAGGGATGATGCAACCAATGCATGGAGGCCAC	459
QY	396	lnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheL	416
DB	460	AAAGAGGTTTCCCGAAAGAGCGGTTTCGGTTCTCCGGCTTGGCTTTTCGAGCATTTTC	519
QY	416	euHisProTyrProLysAspSerLysIleMetLeuAlaLysGlnThrGlyLeuThrA	436
DB	520	TTCAATCCGTATCCTAAAGATTCAGATAAAATCATGCTTTGTAAGCAACACAGGGCTAACAA	579
QY	436	rgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValG	456
DB	580	GGAGCCAGGTATCAAAATTTGGTTTCAATATGCTAGAGTTAGACTATGGNAGCCAATGGTAG	639
QY	456	luGluMetTyrLeuGluGluValLys-----AsnGlnGluGlnAsnSerThrAsnThrS	474
DB	640	AAGA-ATGTACATGGAAGAGTGAAGAAAAACAATCAAGAACAAAAATGGTTTA-----	691
QY	474	erGlyAspAsnLysGluThrAsn	483
DB	692	-----GATCAAAAAAATTATTGAACCTAAT	715

Search completed: July 29, 2005, 18:54:24
Job time : 5786 secs

17c-1000-0111-24

This Page Blank (uspto)